Towards the Design and Synthesis of a p53 Mutant Y220C Rescue Drug

UNIVERSITY
OF SUSSEX

Rhiannon N. Jones

Supervisor: Prof. John Spencer

Submitted to the University of Sussex in part fulfilment of the requirements of the degree of Doctor of Philosophy

August 2017

Declaration

I hereby declare that the work presented in this thesis was carried out at the University of Sussex under the supervision of Professor John Spencer between the dates of September 2013 and July 2017. The work presented in this thesis is my own, unless otherwise stated, and has not been submitted in whole or in part for the award of any other degree.

Signature: ……

Acknowledgments

Firstly, I would like to thank Prof. John Spencer for his guidance, support, encouragement and unwavering enthusiasm. Secondly, I would like to thank our collaborators on this project at the MRC Laboratory of Molecular Biology, Prof. Sir Alan Fersht, Dr. Andreas Joerger (x-ray crystallography), Dr. Matthias Bauer (Biophysical assays and cell viability) and Dr. Matthias Baud for, as well as contributing crucial data to this thesis, all of their helpful insight and guidance as well as unbounding patience. Furthermore, I would like to thank those who contributed to the collaborative research, at Selcia Ltd., Dr. Martin Walker and Dr. Carol Austin, for compound ADMETox determination, and at the University of Sussex, Dr. Cory Ocasio for carrying out the histone H2AX phosphorylation assay.

I would like to thank the University of Sussex and Worldwide Cancer Research (grant no. 14-1002) for funding my PhD.

I would like to thank Dr. Alaa Abdul-Sada at the University of Sussex and the EPSRC UK National Mass Spectrometry Facility for carrying out mass spectrometry, Dr. Graham Tizzard at the UK National Crystallography Service for x-ray crystallography of **3.8**, and Dr. Iain Day for help with NMR.

I would like to thank members of the Spencer group Dr. Adam Close and Dr. Bradley Springett, whom worked on the same project, for their guidance and support, as well as, Dr. Supojjanee Sansook, Dr. Gavin Roffe and Raysa Khan for productive criticism and discussion at the Spencer group meetings as well as their friendship.

I would like to thank Dr. Melvyn Ansell, Katie Duffell, Dr. Jessica Dwyer, Irene Maluenda, Dr. Thomas Moore, Dr. David Neill-Hall, Dr. Oran O'Doherty, and Hayley Rand, for their empathy, friendship and general joviality, as well as many, many, many cups of tea.

Finally, I would like to thank my parents, brother, sisters and my partner Huw, for their support and unconditional love.

Abstract

The DNA damage response is an important barrier to tumorigenesis. Impairment of p53 function is crucial to tumorigenesis by allowing evasion of p53 dependent responses. The mechanisms involve either (i) missense mutations, (ii) partial abrogation of signaling pathways or effector molecules that regulate p53, (iii) epigenetic deregulation.

The tyrosine to cysteine mutation, Y220C, in p53 is found in around 100,000 new cancer cases per annum. This mutation destabilizes the core domain by 4 kcal mol^{-1} and destabilizes p53 under physiological conditions. The large to small mutation results in the fusing of two shallow pockets to create an extended surface cleft that a number of different fragments bind.

The small molecule PK083, 1-(-ethyl-9*H*-carbazol-3-yl)-*N*-methanamine, binds the mutant-specific crevice with a K_D = 150 μM and raised the protein mutant's half-life to over 15 minutes vs. 4 minutes in the absence of the ligand. This presents an ideal starting point towards the design of a p53 rescuing drug.

A library of carbazoles was designed and synthesized, guided by SAR studies, crystallographic information and computational chemistry, with the aim of optimizing the structure toward a more potent PK083 analogue. Affinity gains were achieved by exploitation of direct fluorine-protein interactions between PK9255 (*N*-methyl-1-(9-(2,2,2-trifluoroethyl)-9*H*-carbazol-3-yl)methanamine), and the backbone carbonyls of Leu145 and Trp146 and the thiol of Cys220, resulting in a K_d = 28 μM.

Further affinity gains were achieved through SAR studies targeting the proline-rich subsite II. Chemistry was optimized to allow a diversity-oriented synthesis toward 2,6,9-substituted carbazoles. A small library of PK083 analogues, where the subsite II targeting group was a halogen, ether, ester, amide or heterocycle were synthesized, identifying the heterocyclic compounds as most potent. A scan of heterocyclic compounds was carried out to identify the most potent heterocyclic substitution.

Abbreviations

ABL	Abelson Kinase
ADME	Absorption, distribution, metabolism, excretion
AKT	Protein Kinase B
APAF1	Apoptotic protease activating factor 1
ARC	Apoptosis repressor with caspase recruitment domain
ARF	Alternate reading frame
ATM	Ataxia-telangiectasia mutated
ATP	Adenosine triphosphate
ATR	Ataxia telangiectasia and Rad3 related
Bcl2/6/XL	B-cell lymphoma 2/6/XL
BCR	Breakpoint cluser region protein
BH3	Bcl-2 homology domain 3.
BRCA1/2	Breast cancer susceptibility protein ½
CCSD(T)	Coupled cluster single double (triple)
CdK2/4	Cyclin-dependent kinase-2/3
CETP	Cholesterylester transfer protein
Chk1/2	Checkpoint kinase 1/2
CNS	Central nervous system
CSD	Cambridge structural database
CTD	Carboxyl-terminal regulatory domain
CytC	Cytochrome C
CytP450	Cytochrome P450
DBD	DNA-binding domain
DCM	Dichloromethane
DFT	Density functional theory
DMA	Dimethylacetamide
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
DPP	Dipeptidyl peptidase
Dr5	Death receptor 5
DSBs	Double-stand breaks

DSF	Differential scanning fluorimetry
E3	Ubiquitin ligase
EGFR	Epidermal growth factor receptor
FACT	Facilitates chromatin transcription
FAS	Apoptosis antigen 1
FDA	Food and drug administration
GNM	Gaussian network model
HSA	Human serum albumin
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HER	Human epidermal growth factor receptor
hERG	Human ether-a-go-go gene
HIC1	Hypermethylated in cancer protein 1
HPV	Human papilloma virus
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence
IBMP	2-methoxy-3-isobutylpyrazine
IGF	Insulin-like growth factor-binding protein
INK	Inhibitors of cyclin-dependent kinases
IPMP	2-methoxy-3-isopropylpyrazine
ITC	Isothermal titration calorimetry
JNKs	c-Jun-N-terminal kinases
KSP	Kinesin spindle protein
lc-ms/MS	Liquid chromatography-tandem mass spectrometry
LEF	Local environment of fluorien
LFS	Li-fraumeni syndrome
MALDI-TOF MS	Matrix-assisted laser desortion/ionization time-of-flight mass spectrometry
MAOS	Microwave-assisted organic synthesis
MD	Molecular dynamics
MDM2/X/L	Mouse double minute 2/4/X/L
MPO	Multiparameter optimization
MIDA	N-Methyliminodiacetic acid

MOMP	Mitochondrial outer membrane permeabilization
MP2	Second order Møller-Plesset perturbation theory
MPTP	1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine
mRNA	Messenger ribonucleic acid
mTor	Mechanistic target of rapamycin
MUP	Major urinary protein
Myc11	Myelocytomatosis viral oncogene homolog 11
NADPH	Nicotinamide adenine dinucleotide phosphate
NBS	N-bromosuccinimide
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHC	N-heterocyclic carbenes
NMR	Nuclear magnetic resonance
Noxa	Phorbol-12-myristate-13-acetate-induced protein 1
p14	Protein of 14 KDa
p21	Protein of 21 KDa
p44	Protein of 44 KDa
p53	Protein of 53 KDa
PARP	Poly ADP robose polymerase
PD	Pharmacodynamics
PDB	Protein data bank
PK	Pharmacokinetics
PPR	Proline rich region
pRb	Retinoblastoma protein
PTM	Post translational modification
PUMA	P53 up-regulated modulator of apoptosis
RAS	Reterovirus-associated DNA sequences
RING	Really interesting new gene
RITA	Reactivation of p53 and induction of tumor cell apoptosis
RNA	Ribonucleic acid
ROS	Reactive oxygen species
S100B	S100 calcium-binding protein B
SAH	Stapled α-helical peptide

SAR	Structure-activity relationship
SIRT1/2	Sirtuin 1/2
SLM	Short linear motif
SM	Suzuki-Miyaura
SNP	Single nucleotide polymorphism
TAD1/2	Transactivation domain 1/2
TBAF	Tetrabutylammonium fluoride
TBDMS	*tert*-Butyldimethylsilyl
tBid	Truncated BH_3 interacting-domain death agonist
THF	tetrahydrofuran
TP53	Tumour protein 53
UGT	Uridine 5'-diphospho-glucuronsyltransferase
WAF1	Wild-type activating frangment 1
water LOGSY	Water-ligand observed via gradient spectoscopy
Wip	WASP-interacting protein
wt	Wild-type

Contents

Chapter 1 : Introduction

Cancers are one of the leading causes of death worldwide, accounting for 8.2 million deaths in 2012 (GLOBOCAN 2012 (IARC)). The lifetime risk of developing cancer for those born since 1960 is now 1 in 2.[1] Tumourigenesis, the process by which a normal human cell evolves progressively to a neoplastic state, is often likened to a process not dissimilar to Darwinian evolution, in which a succession of genetic changes confers some type of growth advantage over normal cells. Along the road to malignancy, cancer cells acquire a succession of hallmark capabilities that correspond to important barriers to tumourigenesis (Figure 1.1). These comprise of: sustaining proliferative signalling, evading growth suppressors, avoiding immune destruction, enabling replicative immortality, tumour promoting inflammation, activating invasion and metastasis, inducing angiogenesis, genome instability and mutation, resisting cell death and deregulation of cellular energetics.[2]

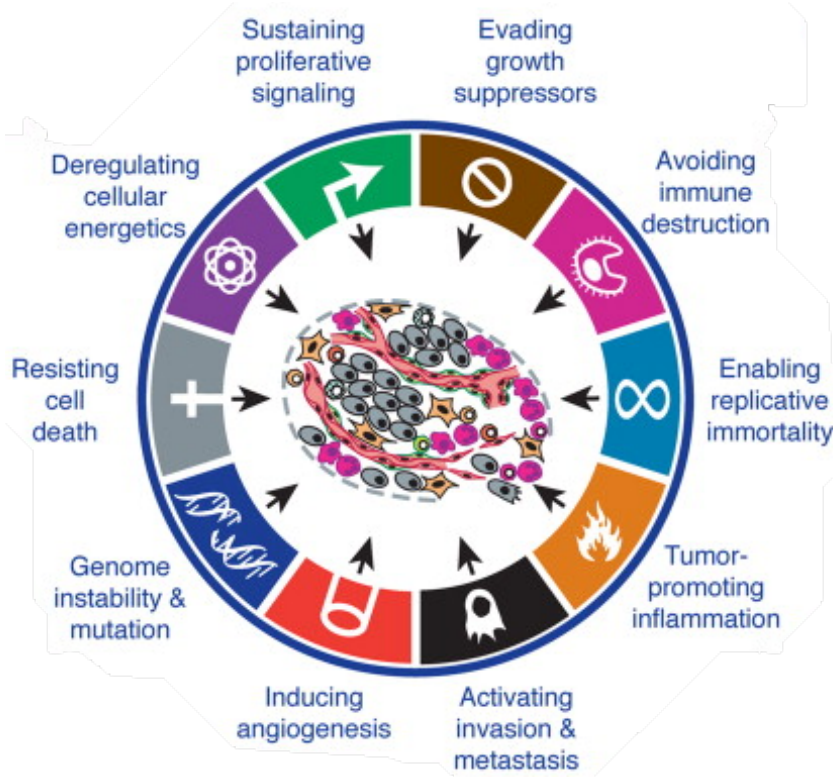


Figure 1.1: The hallmarks of cancer. [2] Reproduced from Cell, 144 (5), D. Hanahan and R. Weinberg, 'The Hallmarks of Cancer: The Next Generation,' 646-674, Copyright (2011), with permission from Elsevier.

1.1 Treating Cancer

Current clinical practice in the treatment of cancer is centred on 8 different approaches: surgery, radiotherapy, chemotherapy, immunotherapy, targeted therapy, hormone therapy, stem cell transplant and precision medicine. The efficacy of each treatment depends on the type of cancer; for example, surgery that works best for cancers which exhibit solid tumours that are contained in one area, however, would not be effective for leukaemia. Therapies are typically used in combination.

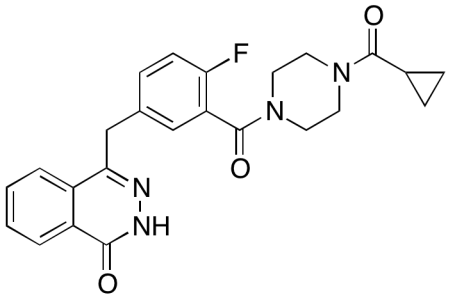
1.1.1 Personalized Medicine

Precision medicine or 'personalized medicine' is an emerging approach to patient care in which an individual's characteristics, including their genetic profile, guide clinical decisions, aiming for a treatment programme tailored to have the maximum efficacy for a specific patient. This treatment course is particularly important for oncology where short-term toxicities and long-term functional implications are associated with surgical and chemo-radiotherapy treatments. Personalized medicine has the power to maximise efficacy of treatment and reduce toxicity by targeting tumour molecular abnormality profile rather than tumour tissue type or anatomical site of origin. The ability to use this approach depends on identification of prognostic and predictive biomarkers that confer the ability to target treatments to patients most likely to benefit.

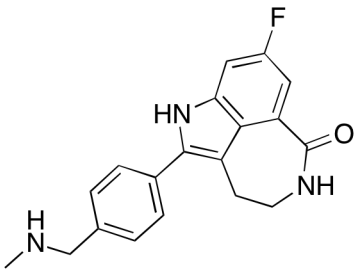
One such example is the identification of cancer predisposition genes BRCA1 or BRCA2, where female carriers of mutations in these genes have a 45-65% chance of developing breast cancer by the age of 70.[3] Screening is available for family members with a strong family history of breast cancer. Current treatment options include mastectomy, oophorectomy or chemical oestrogen deprivation. Although BRCA genes are associated with a high penetrance rate, 30% of carriers will not develop breast cancer, prompting difficult, highly personal decisions.

Cells with BRCA mutations have non-functioning homologous recombination DNA repair mechanisms.[4] However, base-excision DNA repair mechanisms remain functional, rescuing tumour cells from apoptotic death following DNA-damaging cancer therapy. PARP inhibitors prevent base-excision repair from functioning, producing tumour cell death in BRCA deficient cells but not in healthy cells with functional homologous

recombination pathways. The PARP inhibitors Olaparib and Rucaparib have both been granted FDA approval for the treatment of patients with germline mutated BRCA1/2 ovarian cancer (Figure 1.2).



Olaparib



Rucaparib

Figure 1.2: Structure of PARP inhibitors Olaparib and Rucaparib.

Another example is the TP53 gene that encodes the tumour suppressor p53. Evasion of p53-dependent responses is crucial to tumourigenesis with virtually all tumours displaying loss of p53 through either deleterious mutation or partial abrogation of signalling or effector molecules that regulate p53. Consequently, since its discovery in 1979, efforts have been made to target p53 through various avenues toward personalised medicine approaches for cancer therapy.[5] These are summarised in Section 1.7.

1.2 Structural Biology of P53

The p53 tumour suppressor protein acts primarily as a transcription factor that elicits cell cycle arrest, apoptosis or senescence in response to oncogene activation, DNA damage and other stress signals.[6] For this reason, it has been termed 'the guardian of the genome.'[7] Since its discovery, it has also been linked to ageing, metabolism, immunity, embryo implantation, autophagy, angiogenesis and cellular stress. P53's increasingly complex multifaceted role in many signalling pathways is mirrored in its equally complex structural biology. P53 is biologically active as a homo-tetramer, comprising of 4 X 393 amino acid residues (Figure 1.3). At the amino terminus, there is a natively unfolded transactivation domain which can be subdivided into two subdomains TAD1 and TAD2 followed by a proline rich region (PRR). The structured DNA binding domain and tetramerization domain are connected through a flexible linker region. Like the TAD domain, the regulatory domain at the carboxy terminus is intrinsically disordered (CTD).

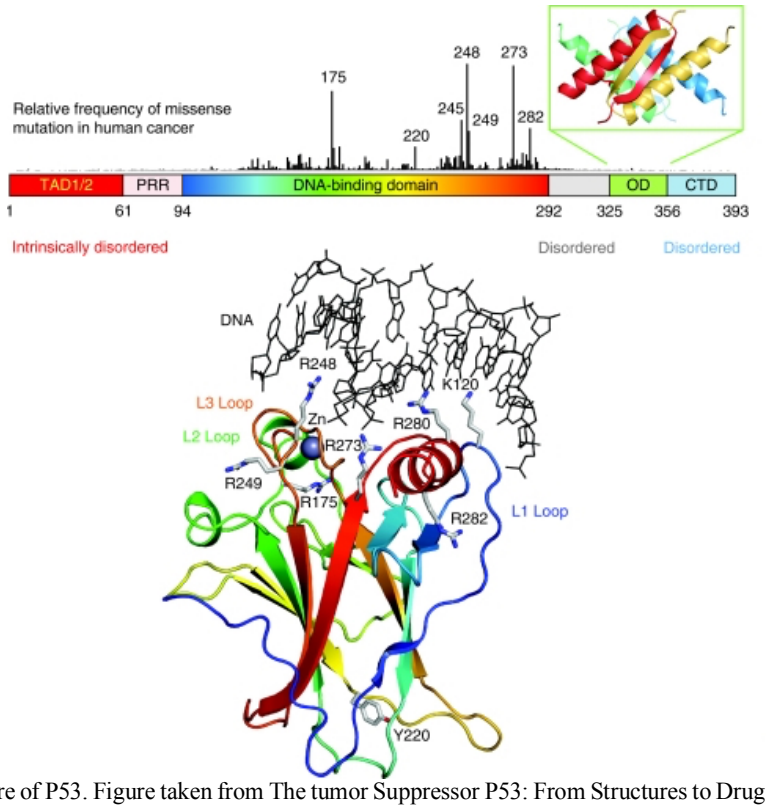


Figure 1.3: Structure of P53. Figure taken from The tumor Suppressor P53: From Structures to Drug Discovery, Joerger et al. [8] Copyright (2010) Cold Spring Harbor Laboratory Press, U.S.A.

1.2.1 The DNA Binding Domain

The DNA binding domain consists of a β-sandwich scaffold as well as a loop-sheet-helix motif and two large loops L2 and L3. The structure of L2-L3 is stabilised by a zinc ion which is tetrahedrally coordinated by Cys176, His179, Cys235 and Cys242.[9–11] The human p53 core domain has low intrinsic thermodynamic stability and rapidly unfolds at body temperature with a half-life of 9 minutes.[12–14] There are several lines of evidence suggesting that this is the result of an adaptive evolutionary process with important implications for protein turnover and binding to partner proteins.[10,15,16] For example, low thermodynamic stability may allow for rapid cycling through folded and unfolded states, which may provide an additional layer of regulation of functionally active cellular protein levels. Additionally, it may also provide a certain element of structural plasticity required to bind many different partner proteins. However, this low intrinsic stability has profound implications on the susceptibility of human p53 to deleterious mutations and cancer development with many mutations reducing the half-life of p53 to minutes or even seconds, thus resulting in loss of functional p53.[12,13]

1.2.2 The Tetramerization Domain

The structure of the individual sub-units of the tetramerization domain consists of a short β-strand followed by an α-helix which is connected through a sharp turn facilitated by Gly344. The p53 folding mechanism proceeds by formation of a dimeric intermediate that is stabilised by antiparallel β-sheet and helix packing interactions.[17] This is followed by dimerization of dimers in a roughly orthogonal fashion via their hydrophobic helix interfaces to form a highly thermodynamically stable tetramer.[18–20] At the centre of this tightly packed interface Leu344 side chains are in direct contact.[21] Leucine 344 is part of a leucine-rich nuclear export signal. The overall effect of this is that the tetramerization domain does, in part, control nuclear levels of p53 through regulation of the oligomerization state. There is evidence that the oligomerization equilibrium of p53 is modulated via an intricate network of accessory proteins that can either have a positive or negative effect. For example, direct binding of apoptosis repressor with caspase recruitment domain (ARC) to the p53 tetramerization domain in the nucleus inhibits tetramerization and promotes nuclear export.[22] In contrast, binding of dimeric 14-3-3 proteins to the p53 carboxy-terminus, which is strengthened by phosphorylation of the latter, enhances tetramerization.[23]

1.2.3 The Transactivation Domain

The natively unfolded TAD domain is responsible for p53's transcriptional activity with two regions of nascent secondary structure.[24–26] Upon binding to the negative regulators murine double minute 2 or X (MDM2, MDMX) through residues 18-25, the TAD1 subdomain has been observed to become fully helical.[27,28] Likewise, regions within TAD2 assume amphipathic α-helices upon binding replication protein A.[29] The interaction properties of the TAD are modulated by posttranslational modification. Out of p53's 9 phosphorylation sites, 7 occur within the binding region of its regulatory proteins (Ser15, Thr18, Ser20, Ser33, Ser37, Ser46 and Thr55) where phosphorylation at different sites modulates affinity for these proteins. For example, phosphorylation of Thr18 significantly reduces the affinity of TAD for the negative regulator MDM2 because of electrostatic repulsion between the negatively charged phosphate group and a negatively charged section on the MDM2 surface.[30–32] Conversely, affinity is significantly enhanced for the transcriptional co-activator p300.[33–36] Therefore, phosphorylation of the TAD is responsible for the modulation of p53 activity in response

to genotoxic stress by modulation of affinity for competing regulatory proteins in the cell cycle by phosphorylation.

1.2.4 The Carboxy-Terminal Regulatory Domain

Like the TAD, posttranslational modifications in the CTD such as: Acetlyation, methylation, ubiquitination, sumolyation and methylation of Lys370, 372, 373, 381, 382 and 386, as well as phosphorylation of Ser366, Ser378, Thr387 and Ser393; are responsible for modulation of p53 function and cellular protein levels.[6,37,38] Additionally, as observed in TAD, intrinsic disorder facilitates binding promiscuity. However, conversely to TAD, in which sections of sequences adopt a specific secondary structure, CTD displays characteristics of a chameleonic sequence where the same sequence can adopt differing secondary structures depending on the structural environment. For example, residues 376-387 form an α-helix in complex with S100B, whereas, parts of the same sequence can adopt a β-sheet secondary structure when in complex with sirtuin 2 (SIRT2).[39,40]

1.3 P53's Cellular Biology

In healthy cells, p53 acts as a 'cellular gatekeeper,' acting to transmit a variety of stress-inducing signals (DNA damage, oncogene activation or hypoxia) to different anti-proliferative cellular responses (apoptosis, cell-cycle arrest, senescence).[41] The process of activation of p53 in response to cellular stress comprises of three steps: stabilization of p53, sequence specific DNA binding and transcriptional activation of target genes.[42] P53 is primarily stabilized through disruption of its interactions with its negative regulator, MDM2. This may be by phosphorylation of specific amino acids in the TAD by various kinases (such as ATM, ATR, DNA-PK, Chk1 and Chk2, see Section 1.2.3), or through antagonism of the p53-MDM2 interaction by the tumour suppressor $p14^{ARF}$.[43,44] P53 then binds DNA in a sequence specific manner.[45] However, several lines of evidence suggest that a significant proportion of p53 is bound to DNA in unstressed cells through the CTD, challenging the stress-induced sequence specific model of DNA binding.[46] One rational is that binding the CTD facilitates DNA binding through the DBD and the search for p53 target sites. Additionally, studies have suggested that basal p53 levels are required for assembly of the p21 pre-initiation complex in unstressed cells.[47] Thus, p53 has been observed to control gene expression in both stressed and unstressed cells.

1.3.1 Cell Cycle Arrest

In mammalian cells, after double-strand breaks by ionizing irradiation, activation of the protein kinase ATM occurs that blocks p53-MDM2 binding through phosphorylation of p53's TAD domain causing a rapid build-up of functional p53.[48] P53-mediated cell-cycle arrest occurs mainly through the transcriptional activation of p21/WAF1.[49,50] P53 binds to two sites 2.4 kb and 1.4 kb upstream of p21.[51] The p21 mRNA is highly induced after p53 activation.[49] P21 binds to the cyclin E/Cdk2 and cyclin D/Cdk4 complexes to cause G_1 cell arrest.[50] This is accomplished through p21 inhibition of Cdk2 and Cdk4 that blocks pRb phosphorylation. This promotes binding of pRb to E2F1, the functional effect of which is transcriptional silencing of E2F1 target genes which are essential to DNA-replication and cell-cycle progression. Experimentally, p21 null mouse embryonic fibroblasts show deficiency in cell-cycle arrest following DNA damage suggesting that p21 is the main transcriptional target of p53 in cell-cycle arrest.[52] However, p21 null cells are not completely defective for G_1 arrest suggesting other p53 target genes.[53]

Furthermore, p53 activation has been found to inhibit G_2/M cell cycle progression. Transcriptional activation of p21, which inhibits cyclin B/Cdc2 causes inhibition of cell cycle progression through mitosis. Additionally, other p53 target genes such as 14-3-3σ may also play a part in blocking G_2/M.[54] P53 repression of cdc28C promoter has also been shown to promote G_2/M arrest after DNA damage.[55]

1.3.2 Senescence

Chronic p53 activation in response to prolonged stress, such as telomere erosion or DNA damage, causes replicative senescence. The ability of p53 to induce cell-cycle arrest is essential for senescence as p21 knockout prevents the induction of senescence by p53.[56] Furthermore, inactivation of p53 can cause senescent cells to re-enter the cell cycle suggesting that cellular senescence is a very stable state of cell cycle arrest but is not irreversible, despite the general belief.[57] There are several mechanisms by which senescence can be sustained. For example, in cells that have persistent p21 gene expression and pRb activation, the level of heterochromatin on E2F1 may be sufficiently high to establish a self-sustaining positive feedback loop to be established.[58,59] Formation of stable heterochromatin may play a role in persistent DNA damage signals, being recognized as DNA DSBs leading to constitutive p53 activation.[60] Although p53

activation facilitates senescence by stress, p53 activation alone is not enough to induce senescence in most cases. Tumour cells in which p53 induction occurs by inhibiting p53/MDM2 only shows signs of reversible cell-cycle arrest.[61] In fact, paradoxically, p53 antagonism of many important proteins for senescence induction, such as pRb, NF-κB, and mTor, may limit senescence. For example, p53 has complex cross-talk through MDM2/pRb binding.[62] Treatment with Nutlin, a p53/MDM2 inhibitor, causes down-regulation of pRb in a p53/MDM2 dependent fashion that may limit the formation of senescent-associated chromatin.[61,63] P53 also inhibits NF-κB, and mTor, both of which are required for the initiation of senescence.[64] Therefore, p53 may facilitate or inhibit senescence depending on the cellular context and pathway cross talk.

1.3.3 Apoptosis

There are several convergent mechanisms by which p53 can induce apoptosis (Figure 1.4), these can generally be separated into intrinsic and extrinsic. Extrinsically, p53-mediated induction of apoptotic machinery proceeds via the induction of death receptors Fas and Dr5 by promoting dimerization, which leads to activation of pro-caspase 8 and ultimately, executioner caspase 3 and 7.[65–67] Intrinsically, p53 induction of apoptosis may be transcription dependent, acting as a transcription factor in the nucleus, or independent, functioning through direct interactions with Bcl2 family proteins in the mitochondria.

Transcription dependent p53-mediated apoptosis proceeds via induction of BH3 domain only pro-apoptotic proteins that causes mitochondrial outer membrane permeabilization (MOMP).[68] MOMP is determined by the balance of pro- and anti-apoptotic proteins. Bax and Bak are activated by interaction with tBid and Bim that insert into the outer membrane and oligomerize to form pores to allow the release of cytochrome C (cyt C). Cyt C binds to ATP and Apaf1 promoting oligomerization to form the apoptosome. The apoptosome recruits and activates procaspase 9 that leads to the activation of executioner caspase 3 and 7. P53 has been found to target multiple stages of the MOMP regulatory mechanism. Bcl2, BclXL and Mcl1 prevent activation of Bax and Bak by sequestering Bim. By signalling for transcription of derepressors Puma and Noxa, p53 can also promote apoptosis by releasing Bax, tBid and Bak from BclXL and Mcl1.

In mammalian cells, a fraction of p53 can be found in the mitochondria preceding MOMP and caspase activation through transcription-dependent mechanisms.[69,70] Transcription-independent p53-mediated apoptosis functions through the direct interactions of p53 with anti-apoptotic and pro-apoptotic proteins. P53 binds BclxL through charge-mediated interactions.[71,72] Additionally, it can interact directly with pro-apoptotic Bak, relieving it from anti-apoptotic Mcl1.[73,74]

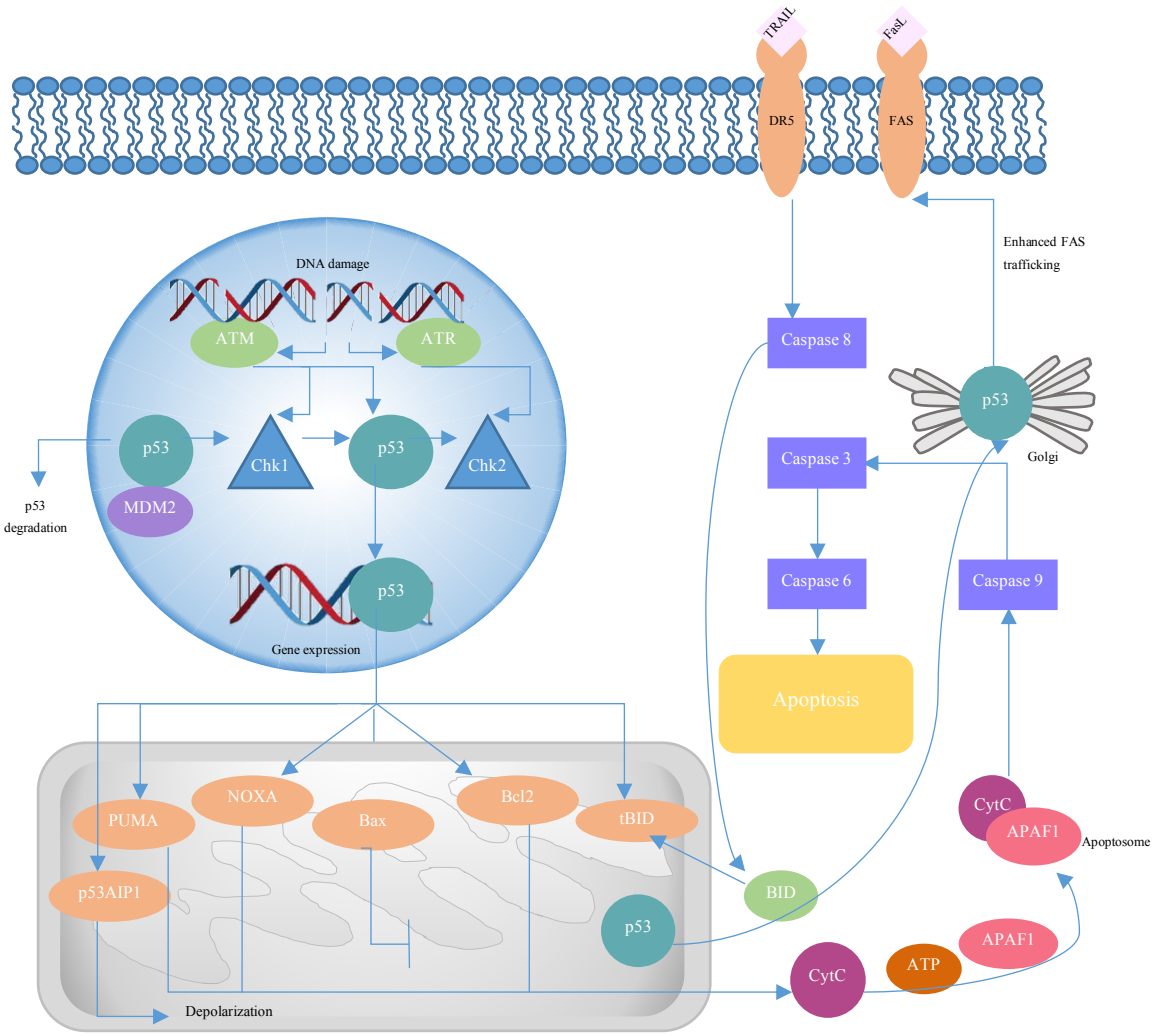


Figure 1.4: Mechanisms by which p53 induces apoptosis.

Uniquely, transcription-dependent and independent p53 activities may cooperate to induce apoptosis in mammalian cells. It has been hypothesised that the structural flexibility of p53 may provide an explanation for this. P53's low thermodynamic stability and high intrinsic disorder may allow the structural plasticity required to bind new

partners without compromising the DNA-binding activity when coupled with high body temperatures in mammalian cells.[13,75]

1.4 Tumour Suppression Function of p53

Numerous studies have investigated how p53 acts as a tumour suppressor, with early emphasis aimed at attributing p53's roles in cell cycle arrest and apoptosis as its key tumour suppression mechanisms. For example, p53 null mice ($p53^{-/-}$) are prone to develop spontaneous thymic lymphomas. Thymoctes ability to undergo apoptosis after irradiation is p53-dependent suggesting that apoptosis is involved in p53's tumour suppression action.[76] Additionally, in choroid plexus brain tumour induced by the SV40 T antigen, mutant antigens defective for p53 binding induced slow-growing tumours that exhibited frequent apoptosis.[77] Wild-type T antigen, a p53 inhibitor, induces rapid-growing tumours. Deletion of p53 in the mutant antigen resulted in rapid growth and reduced apoptosis further suggesting p53 limits tumours by inducing apoptosis.

Several p53 mutants found in human tumours or in the germline of Li-Fraumeni patients have partial transcriptional activity.[78,79] Of these, R175P, E180K and R181L fail to execute apoptosis targets but retain p21 induction suggesting an important role for apoptosis in blocking human tumours. Human derived p53 mutations were introduced into the mouse germline, for example R175P. R172P in humans. R175P mutant mice were largely protected from spontaneous thymic lymphomas which suggests apoptosis is not necessarily needed for tumour suppression.[80] However, this mutation only provided partial protection against Myc-Induced-B-cell lymphoma which suggests that apoptosis is important in this context.[81] Perhaps, the mechanism by which p53 serves as a tumour suppressor differs according to tumour type and cellular context? Furthermore, E180R (E177R in humans) mutant mice, which are deficient for cooperative DNA binding and apoptosis, but not cell-cycle arrest, senescence, and the ability to regulate metabolic target genes, are protected from spontaneous thymic lymphomas and tumours that arose in other organs often remain benign.[82] Hence, apoptosis may not needed to supress thymic lymphoma.

Whilst using p53 missense mutations as a model for examining how p53 acts as a tumour suppressor provides valuable information, to a certain extent, all p53 activities are

compromised to various degrees by mutation. It is unlikely to have a clean separation of apoptosis and cell function.

When apoptosis is blocked in Eμ-Myc mouse by expression of Bcl2, selection pressure to inactivate p53 during lymphoma development was eliminated, showing that in this model for apoptosis, p53-mediated apoptosis is important for blocking tumour progression.[83] Furthermore, Eμ-Myc-induced lymphomas were accelerated by the knockout of Puma which suggests p53 blocks tumour development through apoptosis.[84] Blocking of Myc-induced lymphomas with Bcl2 followed by treatment of established lymphomas with cyclophosphamide caused p53-dependent senescence and tumour regression.[83] Additionally, restoration of normal p53 expression in hepatocellular carcinomas induced by activated H-rasV12 mutant caused tumour regression by senescence.[85] Furthermore, restoration of p53 expression in sarcomas that developed in the absence of p53 lead to senescence and tumour regression.[86,87] All these results point to tumour suppression through p53-mediated cell-cycle arrest, senescence and apoptosis in a context dependent manner. However, recent studies have shown that whilst these mechanisms of action in tumour suppression are undoubtedly important, p53 can supress spontaneous tumour development without inducing apoptosis and cell cycle arrest.

Three p53 acetylation sites are important for the transactivation of classic p53 target genes such as p21 and Puma (K117, K161 and K162). The $p53^{3KR}$ mice showed defects in apoptosis and cell-cycle arrest but were free from early onset tumours.[88] The $p53^{3KR}$ mice retain the ability to regulate target genes involved in energy metabolism and anti-oxidant response. Furthermore, in another triple knockout mouse model, in which p21/Puma/Noxa were knocked out, p53 was found to block spontaneous tumour development showing that the tumour suppressive functions of p53 are not solely dependent on its cell-cycle arrest and apoptotic function.[89] One possible explanation of this enigma is that p53 may induce different target genes to block different stages of tumour development (Figure 1.5).

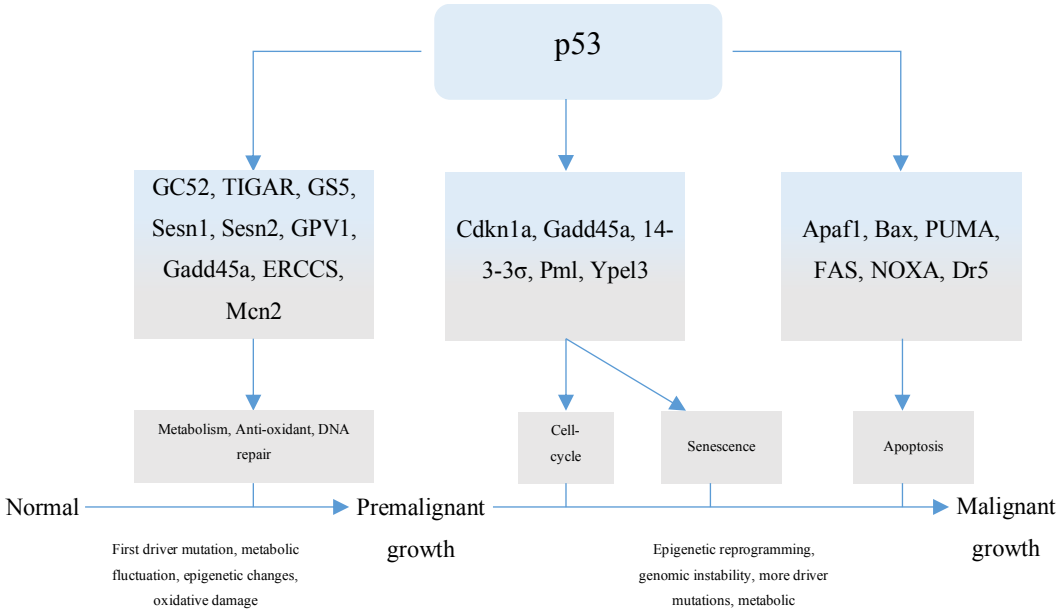


Figure 1.5: Flow diagram of how p53 may induce different target genes to block different stages of tumour development.

1.5 Lessons from Development of P53 Mouse Models: Ageing vs Tumour Suppression

There are numerous mouse models that demonstrate that there exists a delicate balance between the effect that p53 has on tumour suppression and ageing and longevity. P53 null mice develop normally but a subset of the mice exhibit an overgrowth of neural tissue in the region of the mid-brain to cause exencephaly because of defective apoptosis.[90] Furthermore, whilst the mice appear to mature to adulthood normally, the majority seem to develop cancer by 6 months old.[91] Heterozygous $p53^{+/-}$ mice also succumb to cancer, albeit at a later stage than $p53^{-/-}$.[92] Neither mouse model is very useful in studying later onset cancers (which the majority of human cancers are) or ageing and longevity.

$P53^{+/m}$ (mutant p53 allele, 7-11 exons) mice, generated by a spontaneous recombination event which deleted a stretch of DNA upstream of p53, which included exons 1-6 of the p53 coding system, were strikingly resistant to cancer at 18 months compared to $p53^{+/+}$ mice.[93] However, the $p53^{+/m}$ mice has a 20-30% shorter lifespan as well as showing signs of accelerated ageing including tissue atrophy. However, due to the low expression level of the m-allele, no m-derived protein could be detected. Furthermore, in subsequent experiments that characterized the upstream stretch of DNA that was deleted, 24

upstream genes have been identified.[94] It is possible that the accelerated ageing phenotypes and increased cancer resistance in the $p53^{+/m}$ mice could be caused by the absence of one or more of these genes.

Transgenic mice that overexpress p44 (a short, naturally occurring p53 isoform that lacks the main transactivation domain) had a very low incidence of cancer but signs of premature ageing as young as 4 months.[95] By one year of age, most of the mice had died, whereas, non-transgenic mice were healthy. The authors attributed this to p53 hyperactivity which up-regulated the activity of the IGF-signalling pathway which has been linked to drive ageing in diverse organisms such as nematodes, fruit flies and mice.[96]

However, 'super-p53' mice, in which transgenic mice have extra copies of the wt-p53 gene, display up-regulated p53 activities including cancer resistance as well, yet, have a normal lifespan and show no accelerated signs of ageing.[97] Notably, both $p53^{+/m}$ and p44 mice lack the amino-terminal transactivation domain, which is responsible for p53's interactions with MDM2 and hence are not subject to negative regulation. The full-length protein encoded by the super-p53 mice is subject to negative regulation by MDM2.

Mice that express 30% of the normal level of MDM2 ($MDM2^{puro/\Delta 7\text{-}12}$) owing to possessing one hypomorphic and one null allele, showed increased basal levels of p53.[98,99] Upon p53 activation they showed increased expression of p53 target genes and increased apoptosis. $MDM2^{puro/\Delta 7\text{-}12}$ mice were highly resistant to cancer but did not show signs of accelerated ageing. Taken together, it seems that increased p53 levels alone, as observed in the super-p53 mice and the $MDM2^{puro/\Delta 7\text{-}12}$ mouse model, will not automatically lead to development of an ageing phenotype.

Epidermal specific loss of MDM2/p53 signalling in an MDM2-conditional mouse model results in increased p53 levels and increased expression of p53 target genes that are involved in regulating senescence.[100] The mice display hallmarks of early ageing including thinning of epidermal layer, reduced integrity of skin, and widespread and progressive loss of hair. Analysis of the epidermis indicated increased cellular senescence in the follicular bulge region and a reduction of skin epidermal stem cell numbers and functions.

Therefore, it appears that increased p53 levels will only be toxic when the levels exceed the capacity of those proteins which serve to keep p53 activity under tight control. Further research is need to elucidate exact mechanism behind this accelerated ageing phenotype. However, the link between tumour suppression and accelerated ageing hints at potential on-target side-effects of clinical intervention strategies aimed at increasing p53 levels.

1.6 Inactivation of P53

Impairment of p53 function is crucial to tumorigenesis by allowing evasion of p53-dependent responses. There are several mechanisms responsible: through partial abrogation of signalling pathways or effector molecules which regulate p53, by point mutations of p53 itself, and through epigenetic deregulation.

1.6.1 Overexpression of MDM Proteins

A key feature of many tumours is the amplification of MDM genes or altered expression of MDM proteins.[101–108] Tumorigenic increased MDM protein expression is not solely caused by gene amplification, for example, tumours that exhibit high MDM levels without gene amplification are melanoma,[108] Ewing's sarcoma,[109] colon carcinoma,[102] and retinoblastoma.[110,111] Hence, it has been suggested that MDM transcription and translation maybe dependent on tissue type.

As mentioned previously, both MDM2 and MDMX can inhibit p53 transactivation function through engaging its amino-terminal TAD via related N-terminal hydrophobic pockets.[27,110,112] However, key differences between MDM proteins affect their ability to regulate p53 as well as their biochemical functions. MDM2 homo-oligomers exhibit E3 ubiquitin ligase activity that targets p53 for ubiquitin dependent degradation by the proteasome, acting as a crucial negative regulator.[113] MDM2 also inhibits p53 function by preventing its interaction with general transcription machinery. P53 activates the transcription of MDM2, which leads to a decrease in p53 levels, which in turn leads to the transactivation of MDM2 in a crucial negative feedback loop (Figure 1.6).

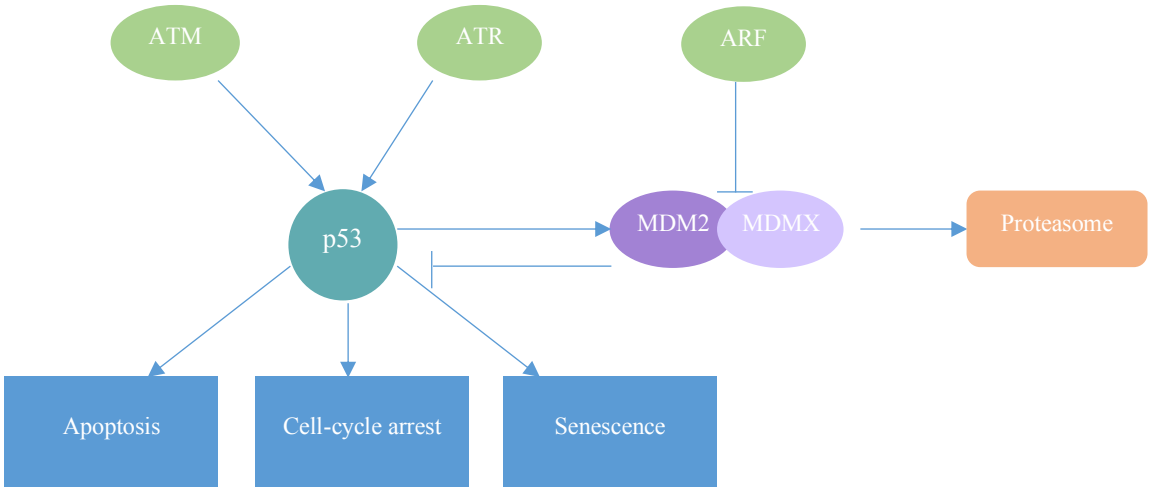


Figure 1.6: Mechanism of regulation of p53 by MDM proteins and ARF.

MDMX is a structural homologue of MDM2 that possesses an 18 amino-acid N-terminal extension.[114,115] MDMX does not homo-oligomerise and has no intrinsic ubiquitin ligase activity. MDMX increases or decreases MDM2 levels depending on MDMX abundance.[116] Hetero-oligomerization of MDM2 and MDMX is crucial during embryonic development.[117,118] Aromatic residues in their RING-proximal domains are required for recruitment of E2-ubiquitin-conjugating enzymes; hence, together they can form more effective inhibitor of p53.[119–121] Although activated p53 was thought to exclusively up-regulate MDM2 through its promoter P2, new evidence has revealed p53-dependent upregulation of MDMX under certain conditions.[115] For example, p53 activation in response to a plethora of p53-activating agents induces the transcription of a novel MDMX mRNA transcript from the MDMX-P2 promoter that is more efficiently translated than that expressed for the MDMX-P1 promoter which controls basal levels of MDM proteins. This encodes a long form of the MDMX protein called MDMX-L that cooperates with MDM2 to promote ubiquitination of p53 hence, p53-induced transcription of MDMX may also play a role in the attenuation of the p53 response.

In addition to MDM transcriptional regulation by p53, levels of MDM proteins are subject to post-transcriptional regulation by microRNA's and post-translational modifications. For example, some p53-induced mRNA's (mir-192, mir-194, and mir-215) contribute to the down-regulation of MDMs.[122] Additionally, multiple kinases are known to phosphorylate both MDM proteins such as DNA-PK, which phosphorylates MDM2 at Ser17,[123] and ABL, which phosphorylates MDMX at Tyr99,[124] both PTMs lead to the dissociation of p53 from MDM proteins. Furthermore, following DNA damage, ATM

and CHK kinases phosphorylate multiple serine residues in or close to the RING domain causing destabilisation of oligomers and hence activating the p53 response.[125–129]

In cancers, these post-transcriptional regulatory mechanisms are most commonly found to stabilise MDM proteins leading to inhibition of p53 and evasion of p53-dependent responses. For example, in multiple myelomas, p53 is wild-type and can induce mir-192, mir-194 and mir-215 which usually down-regulate MDM2 expression.[122] However, epigenetic hypermethylation of the promoter region of all three mRNA's impairs MDM2 down-regulation. Furthermore, some PTM's can lead to the stabilisation of MDM proteins which leads to p53 inhibition. For example, AKT kinases are activated in human cancers and are found to target both MDM proteins through phosphorylation of serine residues (MDM2 Ser166 and Ser186 and MDMX Ser367), the functional effect of which is stabilisation.[130,131] Additionally, using mouse knock-out models, mutation of serine residues usually phosphorylated by ATM and CHK is associated with tumourigenesis.[132,133]

Early transgenic mouse models showed that MDM2 overexpression could induce carcinomas and lymphomas or sarcomas.[134,135] Furthermore, in p53 null mouse models, MDM2 overexpression does not seem to accelerate the onset of tumorigenesis, hence, MDM2's major oncogenic effect must be through p53 inhibition. Conversely, mouse models of MDMX overexpression have yielded conflicting results whereby differences in transgene induction strategies and mouse genetic backgrounds providing likely explanations for conflicting results, indicating a role of tissue specificity in the oncogenic activity of MDMX. However, *in vitro* studies point towards the necessity of the p53-biding domain of MDMX suggesting it exerts its oncogenic role in the same manner as MDM2.[136,137] For example, transduction of normal human and mouse fibroblasts with oncogenic RAS variants leads to senescence which is generally p53-dependent. Co-expression of MDMX leads to transformation.[138,139] For example, melanocyte-specific expression of MDMX and RAS resulted in early onset tumours that were more aggressive.[108] Thus, seeking to antagonize the MDM-p53 protein-protein interaction is an attractive target for drug development. A summary of current strategies is provided in Section 1.7.1.

In addition to transcriptional regulation by p53, MDM2 is under tight surveillance by the nucleolar tumour suppressor protein, $p14^{ARF}$ that is encoded by the INK4b-ARF-INK4a locus.[140,141] Abnormal proliferation in response to oncogene activation signals for the transcription of $p14^{ARF}$ that directly binds MDM2 at a site distinct from the p53 binding site and inhibits ubiquitin ligase function and sequesters MDM2 into the nucleus.[142] Together with p53 and pRb, $p14^{ARF}$ represents one of the most frequently inactivated tumour suppressors in human cancer.[143] Thus, loss of $p14^{ARF}$ can lead to evasion of p53-dependent responses.

In over 50% of tumours that exhibit p53 inactivation, it is found that overexpression of MDM2/MDMX as well as deletion of $p14^{ARF}$ lead to loss of p53 function.

1.6.2 TP53 Mutation

In the other proportion of tumours, p53 is inactivated through mutation. The prevalence of p53 mutation varies significantly according to cancer type and the developmental stage of the tumour (Figure 1.7). In contrast to other tumour suppressor genes that are mainly altered by truncating mutations, most TP53 mutations are missense resulting from single nucleotide substitutions. Other alterations in TP53 include frameshift insertions and deletions, nonsense mutations, silent mutations and other infrequent alterations.[144] TP53 missense mutations found in cancer can broadly be divided into two categories, those which remove essential DNA interaction sites, or contact mutations, and those which perturb the structure of the DBD, structural mutations (Figure 1.8).

Contact mutations arise from nucleotide substitutions at highly mutable CpG dinucleotides, at codons encoding residues that play essential structural and chemical roles in p53's sequence specific DNA binding (contact mutations).[145,146] Mutations of any of the three key arginine residues in the p53 DNA-binding surface, Arg248, Arg273, or Arg280, result in the loss of sequence specific DNA binding and transcriptional activity.[147,148] For example, contact mutations R273C and R273H both remove a guanidinium group that forms important contacts with the DNA backbone at the centre of the p53-DNA interface with little effect on p53 structure or stability.[149–152]

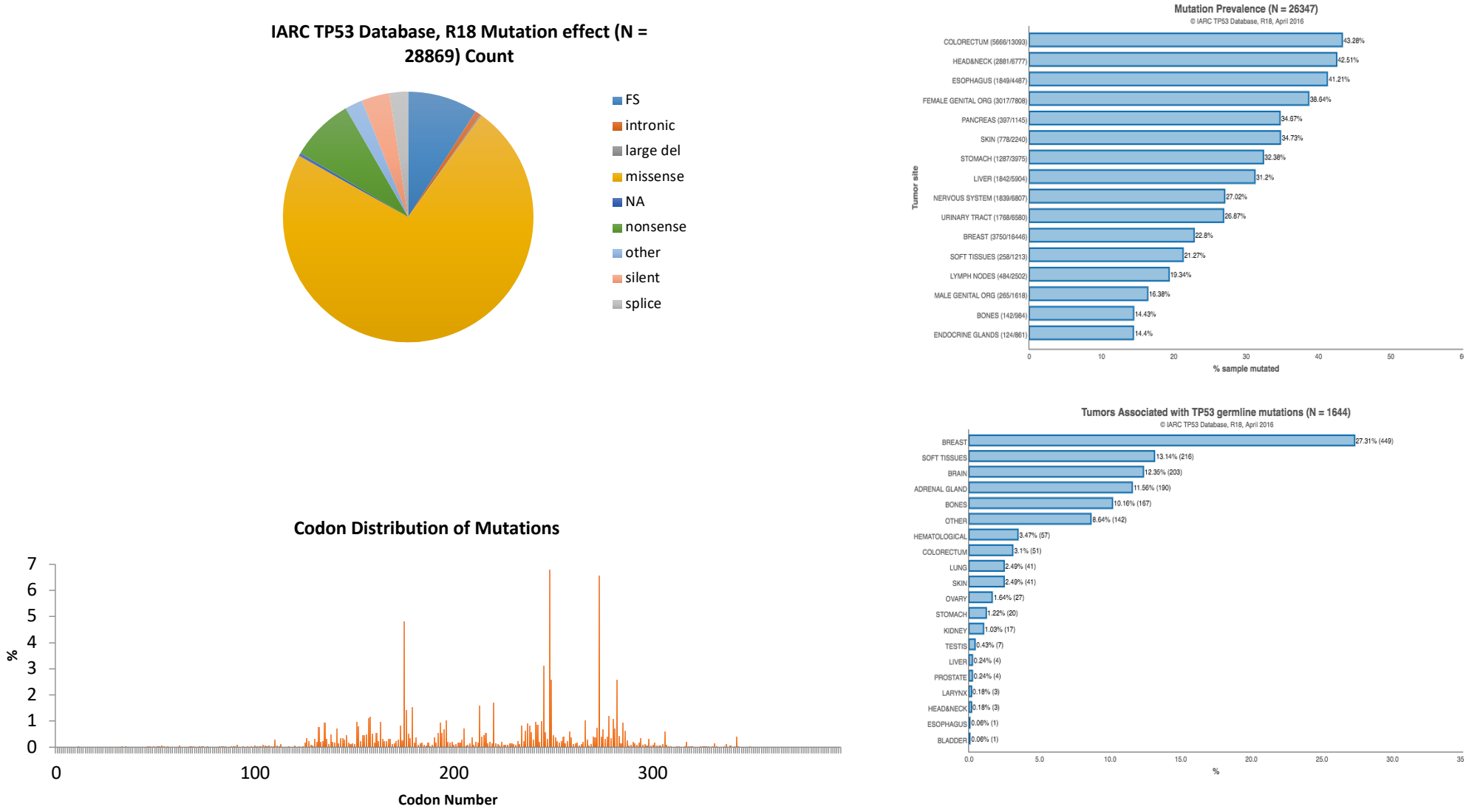


Figure 1.7: Mutation distribution in TP53 and frequency of somatic and germline mutations as a function of associated cancers. Figure generated from data acquired from the IARC p53 database.[144]

Furthermore, the R248W mutation not only removes a key interacting residue but also directly blocks DNA binding by the introduction of a large hydrophobic residue.

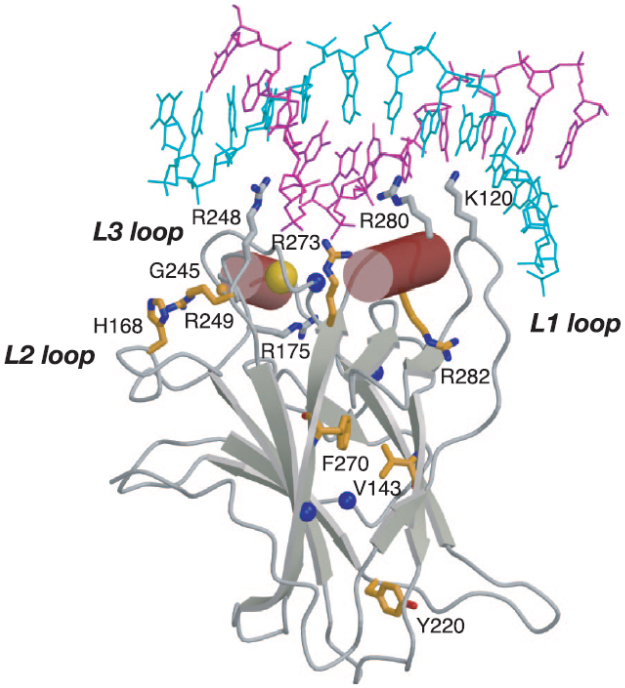


Figure 1.8: Missense mutations in p53 as characterized by X-ray crystallography. Figure taken from Joerger et al.[153] Copyright (2006) National Academy of Sciences, U.S.A.

Structural mutations differ significantly in their destabilizing effect on the p53-DBD depending on the location of the mutation and the type of mutation.[147,154]Large to small residue mutations, such as V143A and F270L are highly destabilising as they create large internal cavities and result in the loss of hydrophobic interactions.[150] In contrast, the small to large oncogenic V157F mutation results in an energetically unfavourable repacking of the hydrophobic core of the DBD around the mutation site.[155] The common feature of these mutations is that they destabilise the core domain by 3-4 kcal mol^{-1}, resulting in lowering the melting temperature of the protein by 5-7 °C.[12,147] As the wild-type core domain is only marginally stable and has a melting point only slightly above physiological temperature, highly destabilized mutants are unfolded under physiological conditions and are no longer functional.[14] However, many structural mutations in the DNA-binding surface are not associated with temperature sensitive phenotypes because, in addition to destabilising the protein, they induce local structural changes that impair DNA binding in the folded state. Disruption of the Arg249-mediated hydrogen bond

network by the R249S mutation perturbs the hairpin conformation of the L3 loop that is crucial for the correct positioning of the DNA-contact residue Arg248.[149,156]

Mutations that result in the loss of zinc are particularly deleterious as they impair DNA binding as the zinc is crucial for the integrity of the L2-L3 region that interacts with the minor groove of DNA response elements.[147] These mutations either effect the zinc ligands, such as C176F, H179R, C238Y and C242S or residues in close proximity to the zinc coordination sphere, such as R175H. [148]

1.6.2.1 Spotlight on Y220C

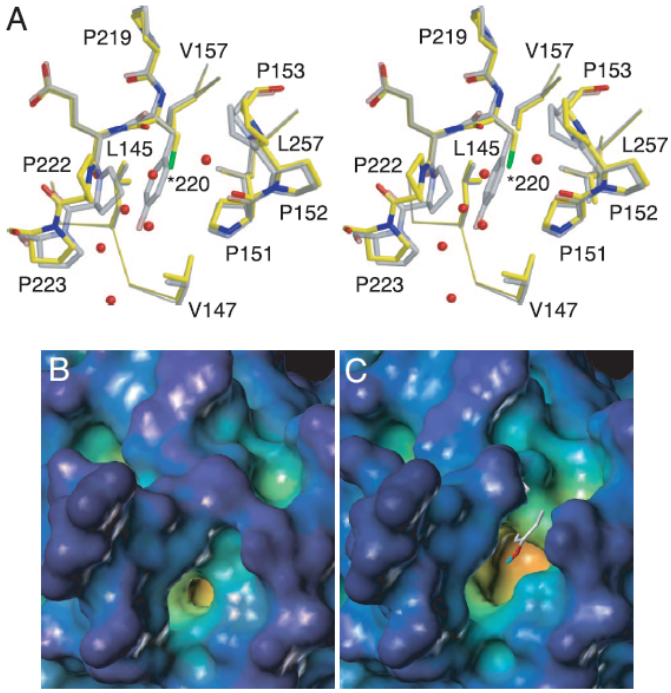


Figure 1.9: Crystal structure of T-p53-Y220C. (A) Stereoview of the mutation site at the periphery of the β-sandwich in T-p53-Y220C superimposed on the structure of T-p53C. (B) Molecular surface of T-p53C around Tyr-220. (C) Molecular surface of T-p53-Y220C (view the same as B). The position of Tyr-220 in T-p53-C is shown as a stick model. Copyright (2006) National Academy of Sciences, U.S.A.

The most common mutation outside the DNA-binding domain is Y220C, which occurs in around 100,000 patients per annum.[154] The missense mutation occurs at the far end of the β-sandwich, at the start of the loop connecting β-strands S7 and S8. It has been found that Y220C destabilises the core domain by 4 kcal mol^{-1} and thus destabilises p53 under physiological conditions.[150] The benzene of Tyr-220 (Y220) forms part of the hydrophobic core of the β-sandwich, with the hydroxyl group pointing towards solvent.

The crystal structure of T-p53C-Y220C, an engineered stable variant of p53 possessing the Y220C mutation, showed that the Y220C mutation creates a solvent accessible cleft but leaves the overall structure of the core intact (Figure 1.9). The structural changes upon mutation lead to the 'fusion' of two shallow clefts to create a crevice, which has its deepest point at the Y220C mutation. The positions of the peripheral hydrophobic side chains are conserved; however, the mutation leads to a loss of hydrophobic interactions.

1.6.2.2 Li-Fraumeni Syndrome

Li-Fraumeni Syndrome (LFS) is a rare inheritable cancer predisposition syndrome which is caused by heterozygous germline mutations in TP53.[157] Individuals with LFS are predisposed to develop a number of cancers including soft tissue and bone sarcomas, breast cancer, CNS tumours, adrenocortical carcinomas and acute leukaemia's as well as a spectrum of other neoplasms that occur less commonly but at higher frequencies and at younger ages compared with the general population.[158–163] It is currently estimated that the mutation carrier rate is 1 in 5,000.[164] LFS is more common in southern Brazil because of the presence of the R337H founder mutation that has a high population prevalence of 0.3%.[165–167] This mutation occurs at exon 10, which codes for the TAD. Its functional effect is to destabilize the TP53 tetramer in a pH sensitive manner.[168] The spectrum of LFS-associated TP53 mutations can be separated into two general categories: gain of function and loss of function. Gain of function mutations are missense alterations in the DNA binding domain which confer a dominant-negative effect on wt-p53 function as well as to enable mutated p53 to acquire additional activities that promotes cancer development. Loss of function mutations are nonsense or frame shift mutations, as well as partial or whole gene deletions.[169]

1.6.3 Protein Aggregation

Probably more important than the thermodynamic instability of destabilized mutants is that p53 aggregates in cells.[170,171] The inactivation of p53 and its mutants *in vitro* proceeds rapidly at temperatures below the melting temperature. The DBD unfolds and then aggregates irreversibly to various morphologies depending on the experimental conditions. [12–14,147,172–174] The aggregate forms faster on thermal denaturation of destabilized mutants.[13] The negative dominance of unstable protein mutants has been suggested to be linked with aggregation by denatured mutant nucleating aggregation of wild type in mixed hybrids.[175,176] There is a nucleation-prone sequence surrounding

Ile254 in the DBD that becomes exposed on unfolding, which, it has been suggested, nucleates the aggregation of p63 and p73 and forms co-aggregates in cell lines.[177] The kinetics of aggregation are unusual, involving an initiation event in which two molecules of p53 unfold extensively and condense to start a chain of events in which amyloid-containing structures are formed.[178,179] A cell-penetrating peptide designed to block the amyloidogenic sequence in the DBD has recently been shown to restore p53 tumour suppressor function in ovarian carcinoma.[180]

1.6.4 Polymorphisms in TP53

Most confirmed TP53 polymorphisms are intronic and display no cancer-related phenotypes. However, theoretically, they may affect p53 function through enhanced mutability due to DNA sequence context, increased cryptic splicing events, altered transcript stability, translations or tissue-specific expression.[181,182]

Out of the 19 confirmed polymorphisms in the TP53 coding sequences, 8 are synonymous. Whilst they do not change the amino-acid sequence or have no effect on protein structure, in theory, modifications in the base sequence could affect protein expression, folding and function as well as provoking new splicing events.[183–185] For example, a synonymous mutation at codon 36 (CCG to CCT) was shown to reduce the ability of p53 to activate apoptosis by lowering the affinity of TP53 RNA for MDM2 (Table 1.1).[183] Taken together with the finding that MDM2 binds to p53 RNA and facilitates its translation, this polymorphism leads to a reduction in functional p53.

The 11 remaining polymorphisms are non-synonymous and result in a change in the amino-acid sequence of p53 which, as for mutation, can affect the ability of p53 to bind to response elements of target genes, alter recognition motifs for PTMs or alter the protein stability and interactions with other proteins.[186–188]

Whilst there is undoubtedly a link between SNPs and an increased cancer risk, attempts at associating specific polymorphisms with an increased risk of specific cancers has often provided contradictory results. For example, links between homozygous R72 carriers and increased susceptibility to cervical cancer by associating its increased degradation by the E6 protein of HPV16 over P72 have largely proved statistically insignificant.[189,190]

Regardless, it is likely that future studies in TP53 SNPs will help provide predictive biomarkers for prevention and early intervention strategies.

Polymorphism	Variation	Effect	Outcome	Reference
Codon 36	CCG – CCT	• Lowers affinity of TP53 RNA for MDM2.	Reduction in functional p53.	[183]
Codon 47	CCG – TCG (P47S)	• Reduction in phosphorylation by proline-directed kinase p38 MAPK • Decreased ability to transactivate PUMA and p53 AIP1	Reduction in apoptotic function through alteration in recognition motifs for PTMS	[191]
Codon 72	CGC – CCC (R72P)	• R72 has greater capacity to interact with MDM2. • Facilitates nuclear export into mitochondria.	R72 more efficient at inducing apoptosis than P72 through enhanced interactions with other proteins.	[192]

Table 1.1: Common polymorphisms in TP53 and their functional effect.

1.6.5 Epigenetic Deregulation

Epigenetic modifications such as DNA methylation are well-known to cause a decrease in the expression of thousands of genes. Methylation of coding gene sequences as well as their promoters abolishes gene transcription until a complex process of demethylation restores the expression. During the process of tumorigenesis DNA methylation pattern changes.[193,194] P53 can be inactivated through deregulation of the DNA methylation of its regulatory genes or through direct epigenetic modification in the P1 basal promoter region of TP53.

As p53 is one of the most commonly mutated genes in human cancer, the effect of methylation on the wild-type allele is of importance. Unlike the majority of tumour suppressor genes that are mutated in human tumours, wt-p53 is seldom transcriptionally silenced through methylation but instead, is lost from the genome through deletions of various degrees on chromosome 17.[195] In contrast, epigenetic deregulation of p53 target genes or mRNAs is more commonly observed and has a direct impact on functional p53 levels. For example, p53 activated the expression of hypermethylated in cancer 1 (HIC1)

which in turn regulates sirtuin 1 (SIRT1) to modulate p53-dpenedent DNA damage responses. HIC1 expression in various cancers is silenced by DNA hypermethylation. Loss of HIC1 lead to up-regulated SIRT1 expression in cancer cells. SIRT1 deactelyated and inactivated p53 resulting in cell survival through evasion of p53-dependent responses.[195,196]

However, cytosine methylation of the basal promoter region of TP53 has been shown to reduce the expression of the reporter gene.[197] Indeed, DNA methylation of TP53 gene or its regulatory sequences have been linked to many cancers including: Ewing's sarcoma,[198] glioblastoma,[199] acute lymphoblastic leukemia,[200] human hepatocellular carcinoma,[201] ovarian cancer,[202] breast cancer[203] and multiple myeloma.[204,205] Conversely, lack of DNA methylation of TP53 has been linked to cutaneous squamous cell carcinomas,[206] sporadic adrenocortical cancers[207] and myelodysplastic syndromes.[208] In all examples, direct or indirect activation of the p53 pathway was observed following treatment with the demethylating agent 5-aza-2'-deoxycitidine, followed by an increase in apoptosis.[200] In breast cancer, telomere shortening was correlated with a different level of hypermethylation of TP53.[209] However, the direct mechanism by which TP53 methylation status correlates with tumourigenesis is yet to be elucidated. Furthermore, although DNA methylation is most commonly linked with decreased gene expression, TP53 mutation has been shown to correlate with the TP53 hypomethylation status of exons 5-8.[210,211]

Whilst targeting epigenetic deregulation provides an attractive target for chemoprevention of tumours, proof-of-principle it is yet to be established. Methylation of TP53 exons 5-8 can be increased in rat liver and colon mucosa by dietary selenomethionine, which may influence the preservation of intact TP53 sequence.[212] It remains to be seen if a similar effect could be achieved in human tumours.

1.7 Targeting p53 For Cancer Therapy

In the 40 years since p53's discovery, numerous drug discovery projects aimed at targeting p53 for cancer therapy have been established. These can be divided into two broad approaches, firstly through targeting p53's modulators and secondly, through targeting mutant p53.

1.7.1 MDM2/MDMX Inhibitors

As discussed previously, MDM2 and MDMX are frequently overexpressed in many tumours which leads to loss of functional p53.[213] Therefore, inhibiting either or both of these protein-protein interactions is a viable strategy for the treatment of cancers that exhibit this hallmark.[214,215]

The small molecule RITA has been suggested to activate the p53 response by blocking the p53-MDM2 interaction by directly binding the p53 N-terminal domain (Figure 1.10).[216] However, NMR studies have failed to detect in vitro binding and there is evidence to suggest that RITA's mode of action may be p53 independent, with RITA-induced cell death being triggered by DNA cross-linking.

Inhibitors of the E3 ligase activity of the MDM2 RING domain have been identified however, it has been suggested that inhibiting E3 ligase activity is not sufficient to fully restore p53 function and would need to be co-administered with an MDMX antagonist to achieve the desired therapeutic effect in cancer.[103,217,218]

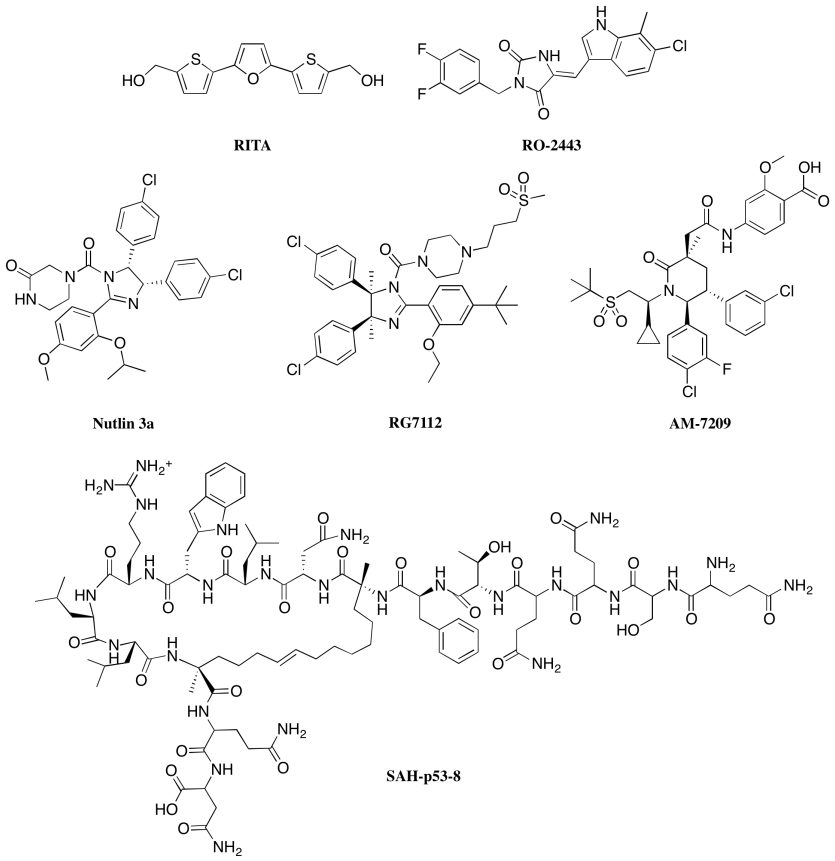


Figure 1.10: Structure of p53-MDM2/X inhibitors.

The first breakthrough in MDM protein inhibitor drug discovery projects was the nutlins, which block the p53 binding site in the N-terminal domain of MDM2 by mimicking three key residues of p53 TAD, Phe19, Trp23 and Leu26 (Figures 1.10 and 1.11A).[219] Since then, numerous classes of small-molecules have been identified to bind to the N-terminal domain of MDM2. The most potent and selective MDM2 inhibitor, AM-7209, binds with a K_D of 38 pM and has shown promising anti-tumour activity *in vivo* (Figures 1.10 and 1.11B).[220] The structure of AM-7209-MDM2 reveals that the intrinsically flexible N-terminal lid region of MDM2 (residues 6-24) is ordered and folds back onto the ligand in an orientation that would interfere with p53 binding (Figure 1.11B).[221] Indeed, there are numerous examples of potent MDM2 inhibitors stabilizing or inducing refolding of the N-terminal lid region. Hence, targeting transient protein states of MDM2 may help in further optimization of MDM2 inhibitors of this type.[222]

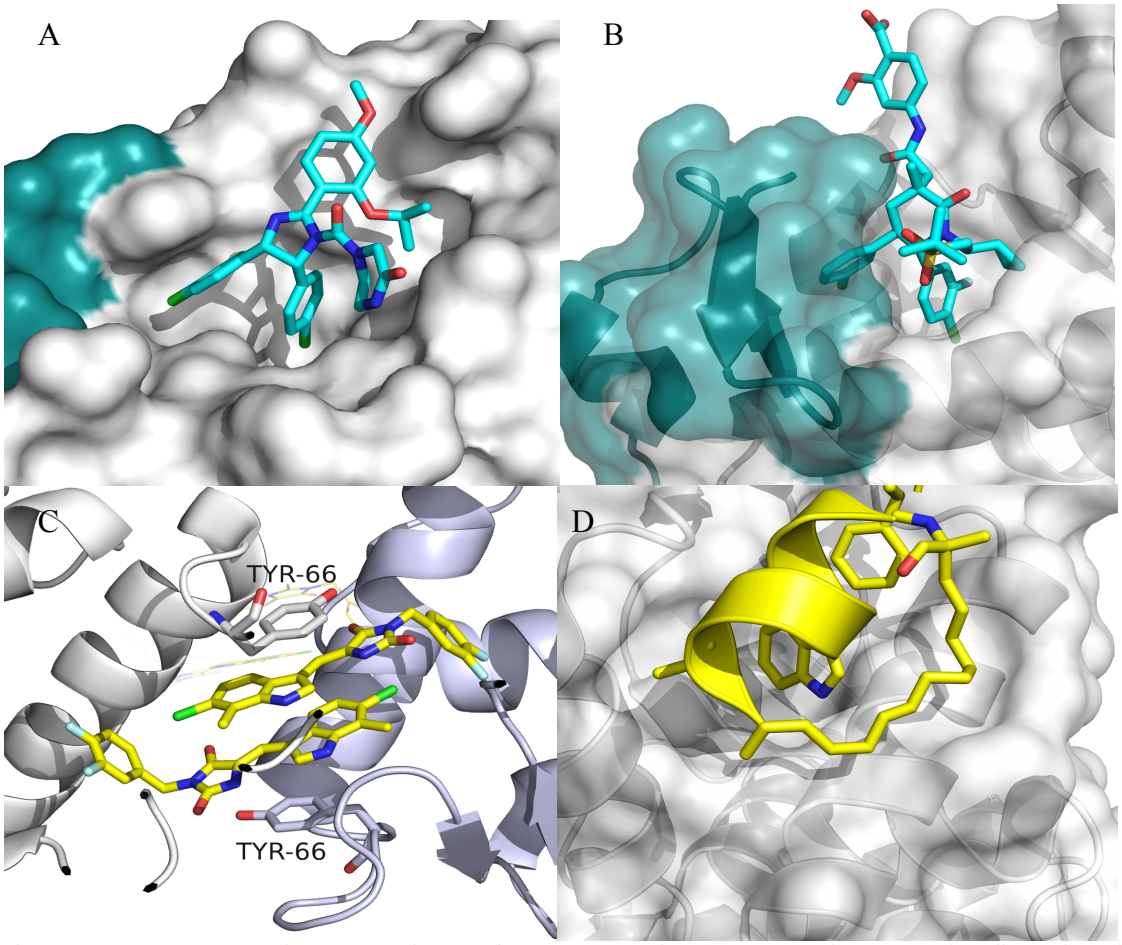


Figure 1.11: X-Ray co-crystal structures of MDM2/MDMX antagonists. (a) Co-crystal structure of Nutlin 3a- bound to MDM2 N-terminal domain, lid region shown in teal (PDB: 4HG7); (b) Co-crystal structure of AM-7209 with MDM2 N-terminal showing the rearrangement of the lid region (teal)(PDB: 4WTZ); (c) RO-2443 binding MDM2 (grey) and MDMX (lilac) (PDB: 3U15); (d) Structure of the α-helical SAH-p53-8 (PDB: 3V3B).

There are several MDM2/X antagonists in clinical trials, Table 1.2 summarizes these.

Class and Specificity	Nature of Compound	Compound	Status	P53	NCT Identifier	Company
Small molecule MDM2 antagonists	Cis-imidazoline	RG7112	Phase I in advanced solid and haematological cancers and liposarcoma (completed)	n/a	NCT00559533	Roche
		RG7112 with cytarabine	Phase I in acute myelogenous leukemia (completed)	n/a	NCT01635296	
		RG7112 with doxorubicin	Phase I in soft tissue sarcoma (completed)	n/a	NCT01605526	
		RO5503781	Phase I in advanced solid cancers (completed)	n/a	NCT01462175	
		RO5503781 with cytarabine	Phase I in acute myelogenous leukemia (active but not recruiting)	n/a	NCT01773408	
		RO5503781 with abiraierone	Phase I/II in advanced prostate cancer (recruiting)	n/a	CRUKE/12/032	
	Spiro-oxindole	SAR405838	Phase I in advanced solid cancers (active but not recruiting)	n/a	NCT01636479	Sanofi-Aventis
		SAR405838 with pimasertib	Phase I in advanced solid cancers (recruiting)	n/a	NCT01985191	
	Imidazothiazole	DS-3032b	Phase I in advanced solid cancers (recruiting)	n/a	NCT01877382	Daiichi Sankyo
	Dihydroisoquinolinone	CGM-097	Phase I in advanced solid tumours (recruiting)	Wtp53	NCT01760525	
	n/a	HDM201	Phase I in advanced solid and haematological cancers (recruiting)	Wtp53	NCT02143635	Novartis
	Piperidines	HDM201 with ribociclib	PhaseIb/II in liposarcoma (recruiting)	Wt053	NCT02343172	Merck
		MK4828 with cytarabine	Phase I in acute myelogenous leukemias (terminated)	n/a	NCT01451437	
	Piperidinone	AMG232	Phase I in advanced solid cancers and multiple myeloma (recruiting)	n/a	NCT01723020	Amgen
		AMG232 with trametinib and dabrafenib	Phase Ib/IIa in metastativ melanoma (recruiting)	n/a	NCT02110355	
	Pyrrolidine	RG7388	Phase 1 in polycythemia vera and essential ihrombocythemia (recruiting)	n/a	NCT02407080	Pegasys
Stapled peptide MDM2/X inibitor	Peptide	ALRN-6924	Phase I in advanced solid cancers (recruiting)	Wtp53	NCT02264613	Aileron

Table 1.2: Overview of current MDM2/X anagonists in clinical trials. Figure adapted from: Clinical Overview of MDM2/X Targeted Therapies, Burgess et al.[223]

In some examples toxicities have been reported. For example, 30% of patients in the clinical trial of RG7112 (Figure 1.10) in liposarcoma experienced grade 4 neutropenia with 15% experiencing prolonged grade 4 thrombocytopenia, however, it is unknown whether this correlates with prior exposure to genotoxic drugs or is entirely RG7112 dependent.[224,225] Furthermore, there are reports of increased p53 mutation following nutlin-3a exposure, causing concern for future development of new cancers in these patients.[226,227]

Theoretically, inhibition of the MDM2-p53 interaction is not specific to cancer cells and normal cells could experience p53 stabilization and increased apoptosis. Furthermore, inhibition of MDM2 could lead to the loss of ability to ubiquitinate other proteins, such as the steroid hormone receptors and pRb, as well as interfering with MDM2's role in DNA repair and modifying chromatin structure.[213] The clinical relevance of these potential long-term on-target toxicities is yet to be elucidated owing to the relative youth of ongoing early phase trials.

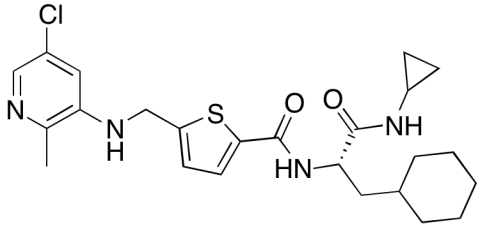
Owing to distinct differences in the Leu26 sub-pocket of MDM2 and MDMX, most developed MDM2 inhibitors are weak binders of MDMX. However, there is growing evidence that for full reactivation of the p53 pathway, both MDM proteins require effective inhibition.[213] Consequently, small-molecule antagonists are being developed which bind both MDM proteins.[228,229] RO-2443 binds both proteins with low nanomolar affinity (Figure 1.10). Binding induces homo- and hetero-dimerization of the MDM2 and MDMX N-terminal domains, which is stabilized by aromatic stacking of two inhibitor molecule and a tyrosine from each of the protein chains (Figure 1.11C).[228] This class of antagonists reactivates the p53 pathway in cancer cell lines overexpressing MDMX that are resistant to conventional MDM2 inhibitors.[228]

In addition to small-molecules, hydrocarbon-stapled α-helical peptides are being developed for treating MDM overexpressed cancers.[230–232] The introduction of a hydrocarbon staple at residues 14-29 of the p53 TAD domain drastically increases the α-helicity of the peptide, and results in a 400-fold affinity increase for MDM2 over the native protein (Figure 1.11).[230] Together with additional mutations added for improved cell permeability, SAH-p53-8 binds MDMX and MDM2 with affinities of 2.3 nM and 55 nM respectively (Figure 1.10, 1.11D).[231,233] SAH-p53-8, and other similar stapled peptides, have been shown to restore p53 function in MDM2/MDMX overexpressing cancer cell lines.[232,233]

1.7.2 Targeting Other p53 Modulators

P53's large network of regulatory pathways provides a large choice of potential targets for cancer therapy. For example, SIR1 downregulates p53 activity by deacetylation of Lys382 and is overexpressed in a number of cancer cell lines, including chronic

myelogenous leukemia stem cells which have wt-p53 status.[234–236] SIR1 inhibition by tenovin-6 in mouse models of chronic myelogenous leukemia led to tumour regression, and was confirmed to selectively kill leukemia stem cells.[236] Alternatively, wild-type p53-induced phosphatase-1 (Wip1) downregulates p53 via the stabilization of MDM2 and modulation of p53-MDM2 binding affinity, and is overexpressed in several cancers.[237,238] GSK2830371 is a small molecule allosteric inhibitor of Wip1 that prevents dephosphorylation of phospho-p53 and induces growth inhibition in both hematopoietic cancer cell lines and Wip1 upregulated breast carcinoma cells with wild-type p53 status (Figure 1.12).[238]



GSK2830371

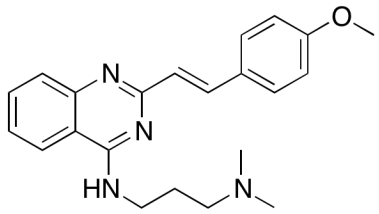
Figure 1.12: Structure of the allosteric Wip1 inhibitor GSK2830371.

1.7.3 Targeting Mutants

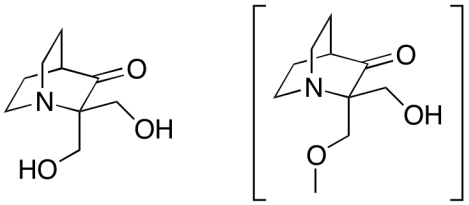
The targeting of p53 mutants with stabilizing small molecules should in theory, stabilize the protein and shift the equilibrium toward the folded state. Therefore, such compounds should restore the active conformation of classical p53 structural mutants. If these mutants have been found to retain wild-type like conformation in their folded states then theoretically, treatment of p53 mutants with stabilizing compounds should lead to reactivation of p53-dependent responses. Stabilizing small molecules can be targeted either against a binding surface shared by the wild-type protein and a subset of mutants or against a mutation-induced lesion on the surface of a particular mutant. Designing molecules for the former is more challenging because of the lack of well-defined binding pockets on the DBD.

1.7.2.1 Alkylating Agents and DNA Intercalators

The first small molecule reported to restore mutant p53 function was CP-31398 (Figure 1.13).[239] However, subsequent analyses in vitro did not detect binding or stabilization of mutant p53 by CP-31398 but instead, showed that the compound acted as a DNA intercalator, which is consistent with cell-based studies demonstrating p53-independent toxicity of the compound.[240,241]



CP-31398



PRIMA-1 **APR-246 (PRIMA-1^{MET})**

Figure 1.13: Structures of CP-31398 and PRIMA-1 (and its methylated form APR-246).

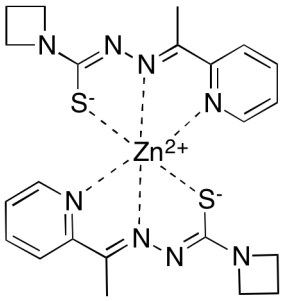
PRIMA-1 upregulates several p53 target genes and induces apoptosis in mutant p53 cancer cell lines potentially via reactivation of a mutant and is currently in Phase Ib/II clinical trials (Figure 1.13).[242,243] Although the anti-cancer properties of PRIMA-1 and related compounds have been extensively documented, its exact mode of action is still being investigated. In cells, PRIMA-1^{MET} is converted to methylene quinuclidinone, which contains a reactive αβ-unsaturated ketone.[243] This structural motif is known to react with nucleophilic cysteines.[244] The alkylation of wild-type and mutant p53 DBD by PRIMA-1 increases the thermostability of the protein, which supports the hypothesis that alkylating agents may play a part in reactivating conformationally unstable mutants. However, the alkylating agent N-ethylmaleimide preferentially modifies Cys182 and Cys277 of the p53 DBD, with further alkylation then triggering cooperative modification of the remaining cysteines in the DBD, unfolding the protein.[244,245] Hence, hyperalkylation interferes with correct folding and function. Furthermore, p53-induced apoptosis stimulated by alkylating agents has been reported for contact mutants R273H and R248Q.[243,246,247] Structurally, it is unlikely that cysteine modification would result in restoration of sequence-specific DNA binding in these mutants. Taken together with the fact that p53 is only one of a myriad of potential alkylating targets in the cell, it seems likely that the mode of action of alkylating agents on mutant p53 may not clear cut.

1.7.2.2 Rescue of Zinc-binding-deficient mutants with metallochaperones

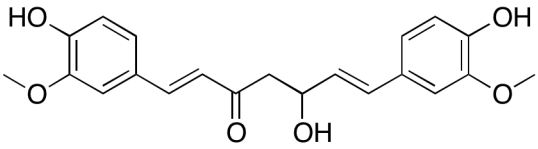
The transcriptional activity of p53 depends on its zinc-binding state; loss of zinc binding results in unfolding, however, an excess of zinc can lead to misfolding and the

aggregation of p53 through binding to non-native sites.[248] Thiosemicarbazone ZMC1 (NSC319726) induces mutant specific growth inhibition and apoptosis in cancer cell lines that possess the zinc-binding deficient R175H mutant (Figure 1.14).[249] ZMC1 acts as a metallocaperone, by binding and transporting zinc ions across the cell membrane as a transition-specific ionophore, increasing intracellular zinc levels to a concentration that suits the repopulation of the distorted zinc-binding site of R175H.[250,251] ZMC1 also exerts a zinc independent effect, potentially involving iron chelation, which affects the redox state of the cell and increases levels of reactive oxygen species (ROS).

Reactivation of mutant p53 has also been reported for a curcumin-based Zn^{2+} complex by restoration of protein conformation and expression of p53 target genes in both the zinc-deficient R175H mutant and the contact mutant R273H (Figure 1.14).[252] However, given that the DNA-contact mutant R273H should be unable to transactivate p53 target genes irrespective of its zinc occupancy, the stabilizing effect of zinc metallochaperone may be related to their DNA-intercalating properties, their reactive Michael acceptor moiety, or both.[253]



ZMC1

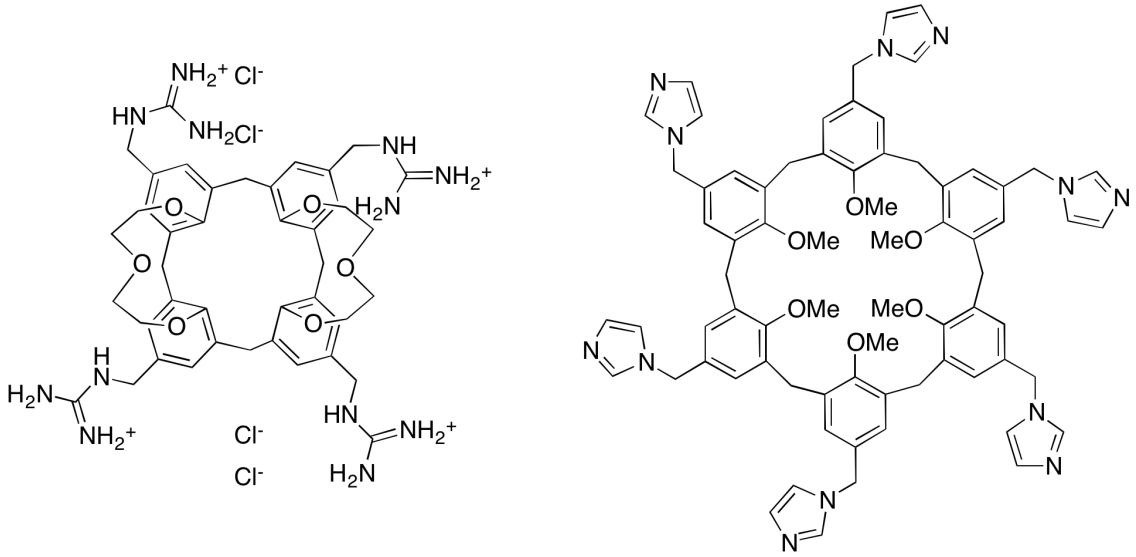


Curcumin

Figure 1.14: Chemical structures of ZMC1 and curcumin.

1.6.2.3 Targeting the tetramerization domain

Most therapeutic concepts are aimed at targeting the most frequently observed mutations, all of which occur in p53's DBD. However, the most frequently inherited mutation, R337H, is found in the tetramerization domain and destabilizes the tetramer in a pH dependent manner.[168] Owing to link between this mutation and Li-Fraumeni Syndrome, which is typified by high penetrance, early onset cancer development, several therapeutic approaches are being perused.



Tetraguanidiniomethylcalix[4]arene **Imidazolecalix[6]arene**

Figure 1.15: Structures of calix arene stabilizers of R337H mutant.

The R337H mutation disrupts a structurally important intermolecular salt bridge between Arg337 and Asp352.[254] The tetrameric state of this mutant can recovered by binding with tetraguanidiniomethylcalix[4]arene (Figure 1.15).[255] Computational and biophysical measurements suggest that two ligand molecules bind to opposite ends of the tetramer and stabilize it through ionic interactions of the guanidinium groups and through hydrophobic interactions of the loops and the lower rim.[255] More recently, an imidazole-calix[6]arene derivative showed improved recovery of the tetrameric state at physiological conditions, and restoration of transcriptional activity of this mutant in cells (Figure 1.15).[256] The tetramerization domain is also being targeted through gene therapy. A replication-defective, TP53-producing adenovirus, Gendicine was granted clinical approval in China but not in the rest of the world.[257] One of the potential limitations of gene therapy is the formation of heterotetramers of reduced activity between wild-type and mutant p53. A potential solution to this is to swap the tetramerizaiton domain for an engineered coil-coil dimerization domain of the breakpoint cluster region protein that forms homo-oligomers without interacting with endogenous BCR or p53.[257] Oligomerization domain swapping may be combined with stabilizing mutations in the DBD or mutations in the transactivation domain that selectively prevent MDM protein binding to create a 'super p53' for future gene therapy.

1.7.2.4 Small Molecule Stabilisers of Y220C

The oncogenic Y220C mutant presents an ideal test case for the design of a mutant-specific drug. The mutation of a tyrosine (Y) to cysteine (C) results in the loss of hydrophobic interaction which destabilises the protein by 4 kcal mol^{-1} and creates a shallow pocket that can be selectively targeted by small molecules.[150] Most importantly, the mutation site is distant from the functional interfaces of the protein, minimizing inhibitory effects from binding that could be the case for DNA-contact mutations in the DBD.

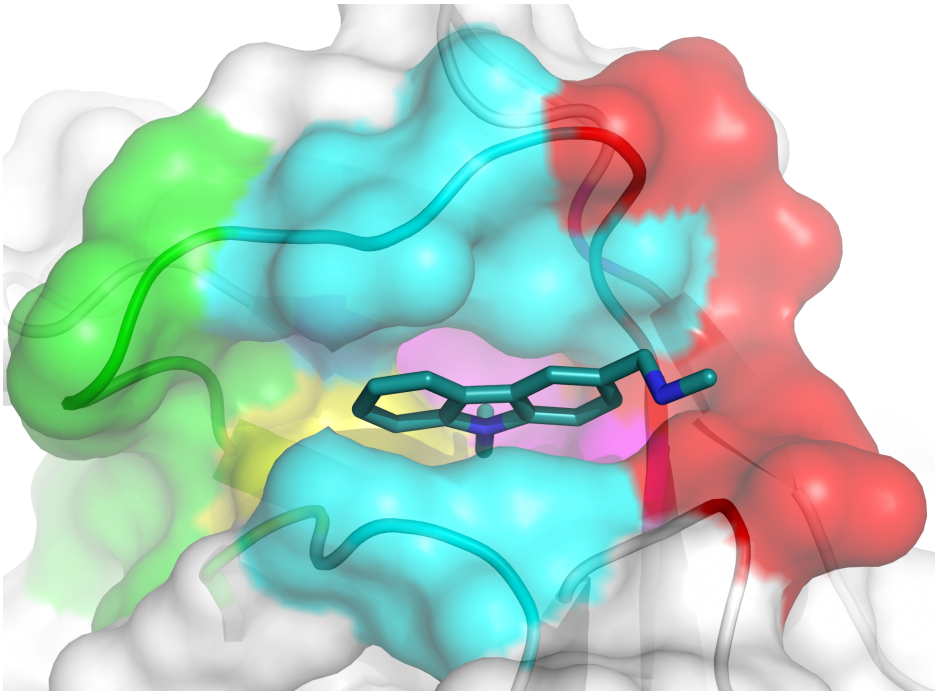


Figure 1.16: PK083 bound to p53-Y220C (PDB: 2VUK). Mutation induced cavity is subdivided into: Subsite I (red), Subsite II (green), central cavity (cyan), Y220C mutation (yellow) and subsite III (pink).

The binding site can be subdivided into a central cavity and three subsites, I, II and III (Figure 1.16). The central cavity consists of several prolines (Pro151, Pro222 and Pro223) which form a hydrophobic interaction surface with Val147 and Thr150. Subsite I is a shallow, solvent exposed subsite that is largely polar. Subsite II is proline rich with Pro153, and several backbone carbonyls (Cys220, Pro151 and Pro152) forming a hydrophobic interaction surface. Subsite III is a transiently open subsite which is essentially modulated by the conformational state of Cys220.[258]

Generally, the compounds which bind Y220C can be classified into two distinct categories, those which bind the closed state of subsite III and those which bind the open state of subsite III as summarized in Figure 1.17.[258]

The lead structure PK7088, is biologically active in cancer cells harbouring the Y220C mutation, increasing the amount of correctly folded mutant protein with wild-type conformation restoring its transcriptional functions and inducing p53-Y220C-dependent apoptosis.[259] Furthermore, aggregation can effectively be inhibited for Y220C because the designed ligands that stabilize the protein also may slow the kinetics of aggregation.[260] These data provide proof of concept that rescuing the function of structural mutants with small-molecule stabilizers is a promising potential cancer therapy strategy. The small

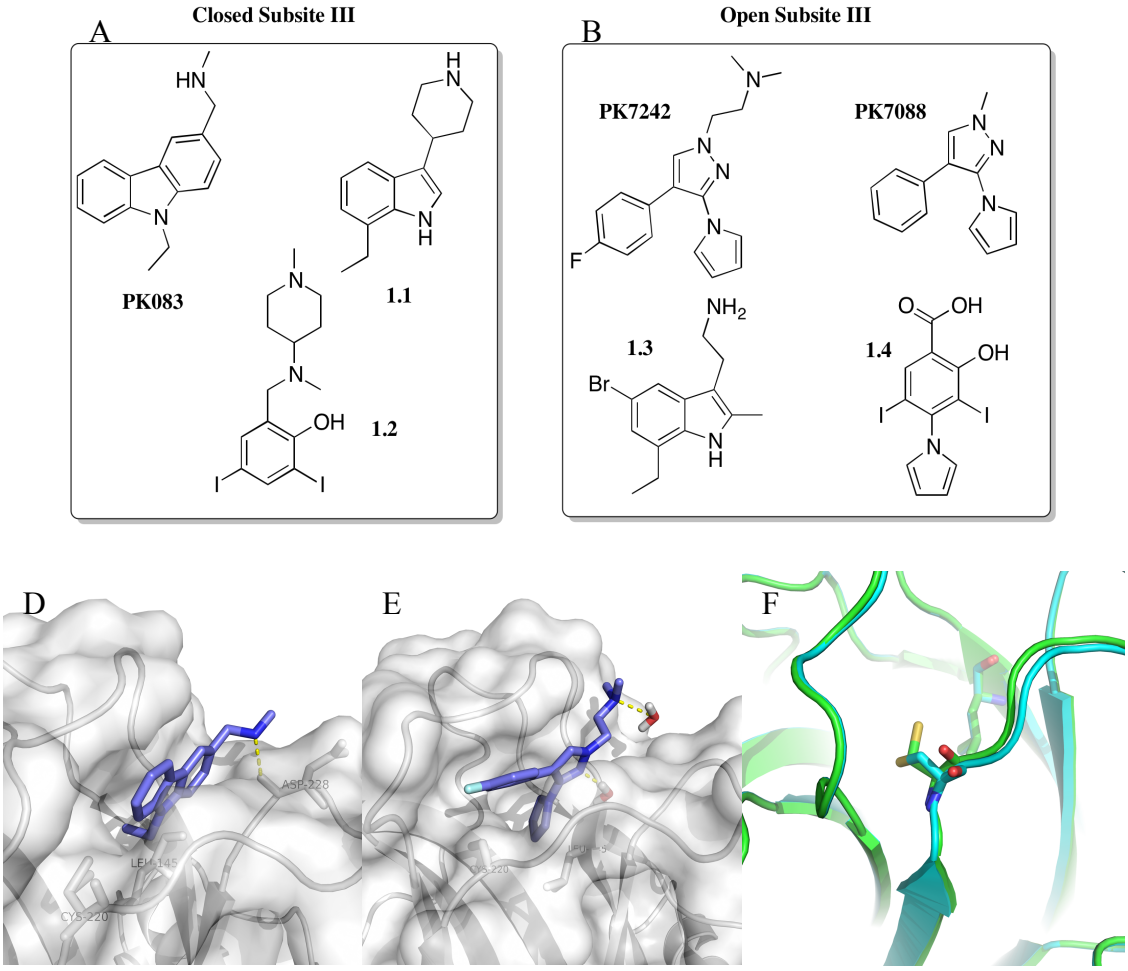


Figure 1.17: Lead compounds which bind Y220C. (A) Structure of PK083 and other classes of fragments which bind Y220C in the closed subsite III state; (b) Structure of PK7242 and other compounds which bind Y220C in the open subsite III state; (d) co-crystal structure of PK083 in Y220C (PDB: 2VUK); (E) Co-crystal structure of PK7242 in Y220C (PDB: 3ZME); (F) Superposition of Y220C in the 'open - green' and 'closed - blue' subsite III states showing the rearrangement of Cys220.

molecule PK083, 1-(9-ethyl-9H-carbazol-3-yl)-N-methylmethanamine, has been identified as one such molecule that binds the mutant-specific crevice. It is found to stabilise the mutant Y220C (K_D = 150 μM) and raise the melting point of p53-Y220C by 2K, as well as increasing the mutant protein's half-life to over 15 minutes compared to

approximately 4 minutes in the absence of the ligand.[261] A co-crystal structure of p53-Y220C with PK083 showed that the central N-Et functionality is buried in the mutant specific hydrophobic cleft, and the NHMe anchor forms a crucial hydrogen bond with the main chain C=O of Asp228 in subsite I (Figure 1.18).

1.8 Thesis Aims

The moderate affinity coupled with the low molecular weight of PK083 presents an ideal starting point for fragment based drug discovery towards the synthesis of a small molecule p53-Y220C stabilizing drug.

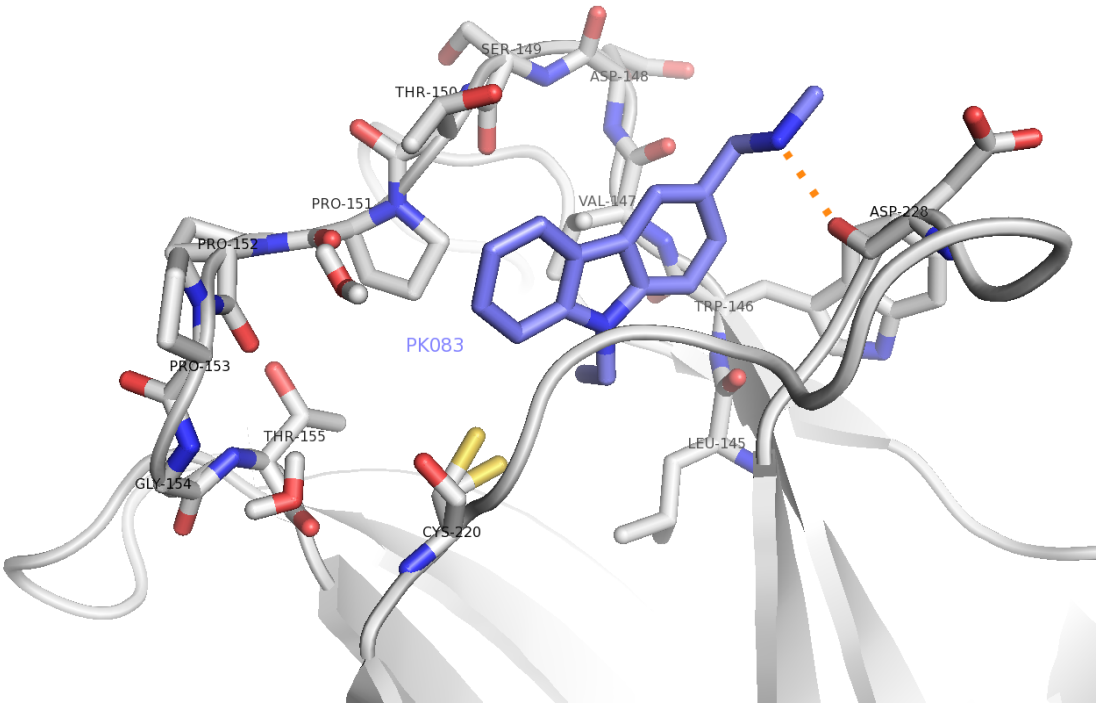


Figure 1.18: Co-crystal structure of PK083 in Y220C showing hydrogen bond between benzylic amine and Asp228 backone C=O, shown in dashed orange line. (PDB: 2VUK).

From examination of PK083 bound to p53-Y220C (Figure 1.18), key points on the carbazole scaffold were identified through which to grow the fragment to target key sub-pockets and/or residues with the aim of enhancing affinity toward p53-Y220C (Figure 1.19).

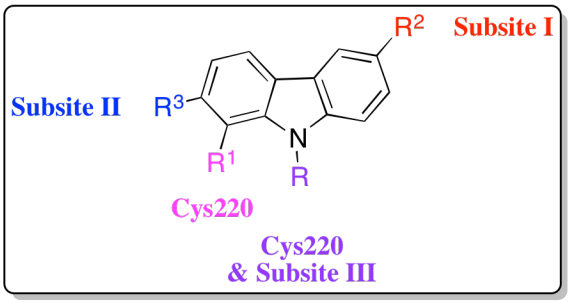


Figure 1.19: Structure of a generic carbazole with key growth vectors highlighted.

This body of work seeks to develop chemistry towards compounds with substitution patterns described in Figure 1.19 and to use the established chemistry to introduce and optimize PK083 towards p53-Y220C.

1.9 Subsite I: Existing SAR

Optimization of R^2 through extension of the amine side chain seeking to probe for potential interactions with the negatively charged Asp228 side chain that at present, the existing anchor is too short to reach is described in Figure 1.20 and Table 1.3.* The first approach was to use rigid aliphatic amines or aromatic amines that, due to the restricted rotation, would not compromise the existing entropically favoured binding mode. Unfortunately, none of these showed a marked increase in Y220C binding affinity. In fact, in most cases, they were worse stabilizers.

The extension of the amine side chain to include acyclic aliphatic side chains largely showed similar binding and stabilisation as for PK083. PK211 possesses an extra hydroxyl group that could interact favourably with Asp228 but does not surpass the affinity of PK083. A co-crystal structure of PK211 with p53-Y220C has been solved. The benzylic amine forms a hydrogen bond between the amine N-H and the C=O of Asp 228. The hydroxyl groups forms an extra hydrogen bond between the –OH group and the Asp228 side-chain. In all known crystal structures, Asp228 points towards bulk solvent so the desolvation penalty is likely to be large. Additionally, the added flexibility of the side chain is bound to be detrimental to binding entropy.

* Unpublished results, R.Wilcken.

Subsite I also possess a tryptophan (Trp146). Attempts were made to target this residue through π-stacking interactions. PK214 exhibits the largest affinity increase. Its side chain features an imidazole ring that is far enough away from the central carbazole scaffold to interact with Trp146. Substitution of the imidazole with a mercaptomethyltetrazole shows a decrease in affinity and results in a less potent stabiliser than in PK083.

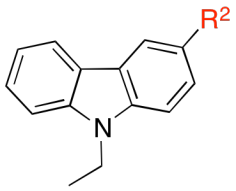


Figure 1.20: Generic structure of carbazoles for targeting Asp228 and Trp146

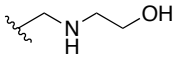
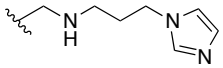
Compound ID	R^2	DSF T_m [250 µM] K	K_D NMR (µM)	K_D ITC (µM)
PK211		0.58	127	199
PK214		0.90	80	98

Table 1.3: SAR of previously tested carbazoles

The failure to make any marked improvements on affinity by targeting subsite I through functionalization of R^2 despite over ~30 compounds synthesized, led to prioritisation of targeting the mutant specific Cys220 mutation through functionalization of R and R^1 (Figure 1.19).

Chapter 2 : Strategies in Improving PK083: Limits of the Cys220 Sub-pocket

The compounds described in Section 1.9 failed to exhibit a meaningful SAR, however, most of the compounds designed and tested were aimed at enhancing potency through targeting Asp228 and Trp146 via functionalization of R^2. Notably, simpler compounds, which seek to establish the effect on potency arising from variation of the methylation state of the benzylic amine at R^2 (e.g. NH_2, NHMe, NMe_2), have not been described.

Following the lack of a meaningful SAR in targeting subsite I for affinity improvements, the mutant specific cavity in Y220C was targeted. Growth at vector R allows targeting of the mutant specific Cys220 hydrophobic subpocket through simple alkylation chemistry, which should confer mutant selectivity and increased affinity (Figure 2.1), or through functionalization of R^1, where R^1 is small, hydrophobic and targets Cys220.

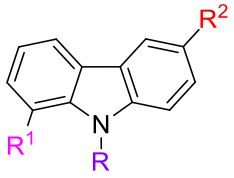


Figure 2.1: Generic structure of synthesised compounds in Chapter 2.

2.1 Methods of Targeting Cysteine Residues

Targeting of non-catalytic cysteine residues is gaining increasing attention from drug discovery scientists owing to the identification of cysteine mutations in several oncogenic proteins.[262] By far the most powerful method of targeting cysteines is via the design of covalent modification compounds, or 'irreversible' inhibitors. Whilst this is the most powerful method of gaining substantial affinity increases, off-target toxicities are commonly reported which are related to lack of specificity.[263–267] Paradoxically, some acrylamide-based modification compounds have been reported to reversibly bind cysteine residues.[268] By tuning the reactivity of the Michael acceptor, by incorporation of a nitrile group in the 3-position, the 1,4-conjugate addition not only seems to be energetically favourable but rapidly reversible. Cysteine also makes an attractive target for non-covalent interactions such as halogen bonding. The relative efficiency of each method is discussed below.

2.1.1 Irreversible Covalent Modification

The design of covalent modification compounds is conceptually very attractive, however, it is very hard to achieve the right balance between reactivity and selectivity. Most covalent inhibitors need to incorporate a highly electrophilic reactive species, such as an α,β-unsaturated ketone (Figure 2.2).[269] However, incorporation of such highly reactive functional groups often leads to toxicities associated with non-specificity, such as alkylation of other macromolecules *in vivo* leading to formation of protein adducts that have been linked to hepatotoxicity,[265] mutagenicity,[267,270] carcinogenicity[271] and immunogenicity.[271–274] Prediction of these off-target toxicities is very difficult, hence, there has been a strong historical bias against the rational development of irreversible inhibitors.[275]

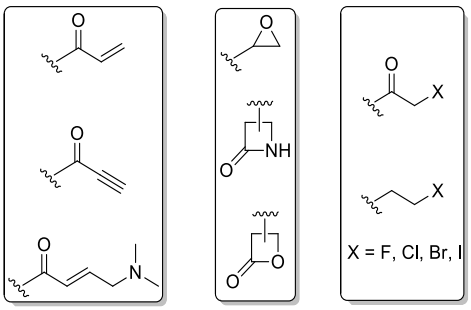


Figure 2.2: Generic structures of common covalent modification warheads.

Pro	Con
Increased biochemical efficiency.	Off-target toxicity.
Lower doses.	Potential immunogenicity could cause allergic reaction or drug hypersensitivity.
Decreased sensitivity to pharmacokinetics (PK).	
Potential to avoid resistance mechanisms.	
Longer duration of action.	

Table 2.1: Pros and cons of irreversible inhibitors.

The potential pros and cons of irreversible inhibition are presented in Table 2.1.[276] Particularly pertinent to oncology is the potential to avoid resistance mechanisms. Inhibitors of human epidermal growth factor receptor tyrosine kinases EGFR and HER-2, Gefitinib and Erlotinib, are not active against the EGFR mutant T790M whereas the irreversible inhibitor Neratinib retains low nanomolar affinity (Figure 2.3).[277] The T790M mutation results in an increase in affinity of EGFR for ATP of an order of magnitude. For inhibitors, Gefitinib and Erlotinib, which both have an ATP-competitive mechanism of action this is deleterious, however, Neratinib does not compete with ATP as its mode of

action involves irreversible inhibition by reaction between its α,β-unsaturated ketone moeity and Cys797, and hence retains activity (Figure 2.3).

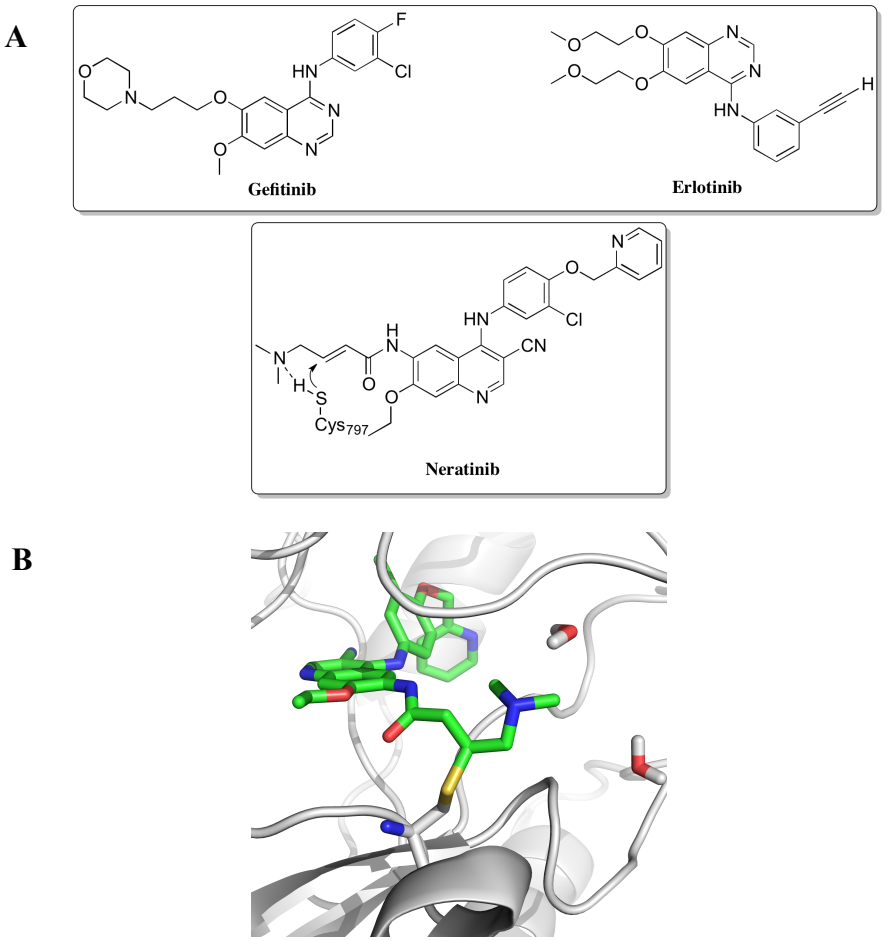
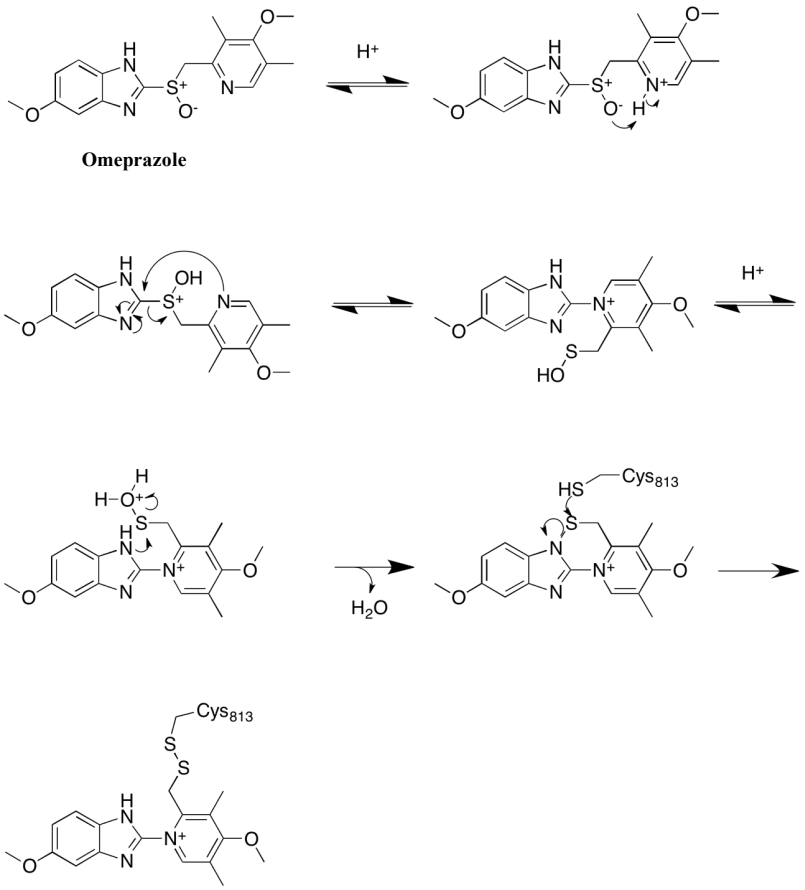


Figure 2.3: Human EGFR tyrosine kinase inhibitors. (a) Structure of reversible inhibitors Gefitinib and Erlotinib and irreversible inhibitor Neratinib. (b) Neratinib bound to EGFR-T790M (PDB: 2JIV). Colour Scheme: Neratinib, carbons shown in green, EGFR-T790M, carbons shown in grey.

In the 1970s much effort was concentrated on the design of 'suicide substrates' that contain masked reactive groups that could be activated according to the drug's mechanism of action seeking to minimize potential off-target toxicities.[278–280] For example, Omeprazole is a covalent modification pro-drug which targets gastric H^+/K^+-ATPase, the enzyme responsible for proton transport as the final step in gastric acid secretion.[281] It is converted under the stomach's acidic conditions into a tetracyclic sulfonamide intermediate that binds covalently to cysteine residues in H^+/K^+-ATPase to form disulfide adducts (Scheme 2.1).[282–284]

Such drugs are extremely challenging to rationally design and are often discovered serendipitously and hence, do not offer a solution to mitigating risk.



Scheme 2.1: Mechanism of action of the prodrug Omeprazole in a highly acidic environment.

Recently, advances in chemical proteomics has led to the development of tool compounds that possess clickable tags, such as alkynes or azides, that allow the evaluation of target selectivity of covalent inhibitors *in vivo* within live cells and organisms.[285–288] In the case that incorporation of an alkyne into a covalent inhibitor does not disrupt binding interactions or affect cell membrane permeability, click chemistry activity-based protein profiling can be used to screen for selectivity. The alkyne analogue of a covalent modifier can be administered to mice and the tissue area of interest can be harvested and subjected to click chemistry conjugation with the desired reporter tag (e.g. rhodamine or biotin) for identification of target proteins.[289,290] The development of activity-based probes may provide a method of de-risking covalent modification compounds.

2.1.2 Reactivity of p53's cysteines towards covalent modification

P53 has 10 cysteine residues (11 in Y220C), most of which are found in the DNA-binding core domain (Figure 2.4). Kaar *et al.* discovered fragments **2.1** and **2.2** bind T-p53-Y220C, raising the melting point by 2 °C (Figure 2.5).[291] Confirmation of the hits by

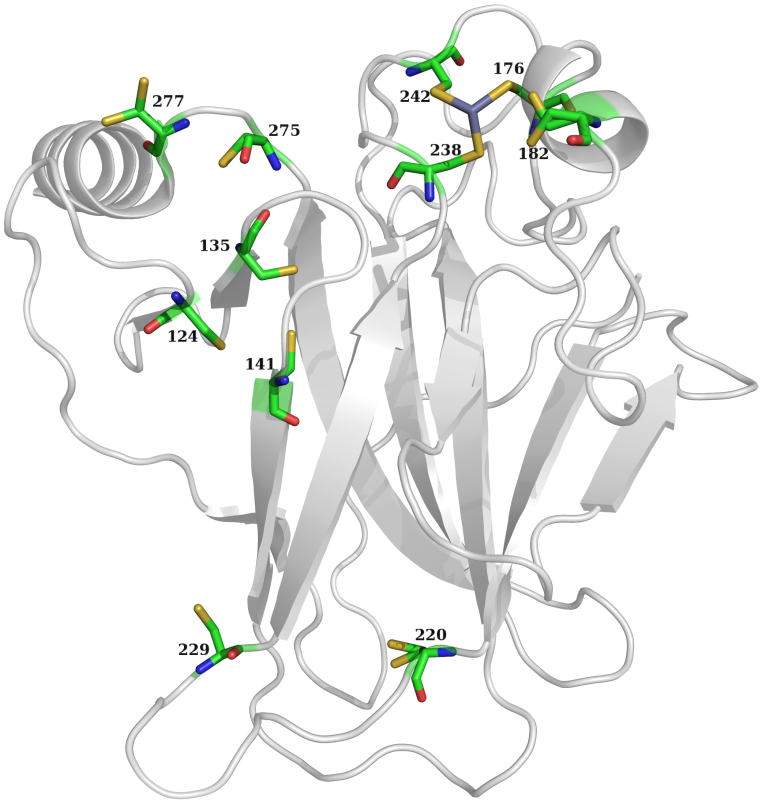


Figure 2.4: p53-Y220C x-ray structure (PDB: 2J1X) with cysteines highlighted.

$^{15}N/^{1}H$ HSQC indicated a shift in residues L114, H115, T123, Q136, C277, R282 and R280, at a site remote from the mutation induced pocket close to the DNA binding domain suggesting a mutation-independent mechanism of action.[245] The effect of **2.1** and **2.2** on other highly destabilised mutants such as R249S, G245S, R282W and R175H using thermal denaturation scanning fluorimetry was determined and indicated that the melting point of each increased (T_m = 1.0 – 3.6 °C) with increasing concentration of **2.1** or **2.2** confirming a mutation independent mechanism. Crystallographic studies detected the alkylation of several cysteine residues indicating various degrees of covalent modification by addition of fragment **2.1** and **2.2** to the α,β-unsaturated double bond.

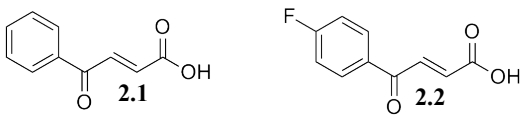


Figure 2.5: Structure of fragments **2.1** and **2.2**.

The reactivity of the 11 cysteine residues in p53-Y220C was assessed by analysing tryptic digests of the modified proteins via data-dependent LC-MS/MS and MALDI-TOF MS at compound to protein ratios of 20:1, 50:1 and 100:1. Mass spectrometric analysis of tryptic fragments of T-p53-Y220C modified by a molar ratio of 20:1 found C124 and C141 to be alkylated indicating they were the most reactive. Modification at a ratio of 50:1 found, in addition to C124 and C141, C135, C182 and C277 to be modified. Modification of C176 and C275 was only observed at the highest compound to protein ratio indicating that they were the least reactive. Modification of C275 and C277 seem to be linked as only one of these sites, not both could be modified under these conditions.

Residue	Solvent-accessible surface (Å^2)	pK_a
Cys124	10.6	13.8
Cys135	0.1	39.1
Cys141	0.1	24.5
Cys176	3.9	10.1
Cys182	21.9	11.7
Cys220	9.7	11.8
Cys229	14.9	12.7
Cys238	0.2	24.6
Cys242	9.3	14.5
Cys275	5.1	20.0
Cys277	19.4	8.3

Table 2.2: Predicted accessibility and pKa of each cysteine in p53-Y220C according to the local protein environment. Solvent accessibility was calculated using the ProtSA server. Theoretical pK_a values of cysteing were calculated using the H++ server, which employs a standard continuum electrostatic methodology. Figure adapted from Kaar *et al*.[245]

The reactivity of cysteines in proteins is modulated by their accessibility, and their nucleophilicity. Intuitively, the reactivity of a cysteine residue towards covalent modification by a drug is greatly affected by the location of the cysteine in the global protein structure. For a reaction to happen, the targeted cysteine should be located at the solvent accessible surface. Additionally, the formation of disulfide bonds between cysteine residues within proteins prevents the targeting of cysteines towards covalent modification. The nucleophilicity of a cysteine residue is greatly affected by the neighbouring residues. Both factors can be quantified using predictions of the solvent-accessible surface and pK_a of the cysteine in its particular environment (Table 2.2). Residues C176, C238 and C242 are relatively inert as they are protected against modification by co-ordination to zinc with modifications only occurring at high concentrations. Surprisingly, the buried residues C135 and C141 were quite reactive.

Modification of C124 could facilitate local rearrangement allowing access and subsequent alkylation of C135 and C141. C182 and C277 are highly solvent accessible and hence are more reactive than C275 which is partly buried. One of the biggest surprises is that C220 is not modified at all. However, fragment **2.1** is polar and the mutation pocket is very hydrophobic and hence fragment **2.1** would struggle to reach C220. The design of a more hydrophobic compound, based on PK083 bearing a Michael acceptor may alkylate C220.[261] However, the development of a covalent modification PK083 analogue should be accompanied by $^{15}N/^{1}H$ HSQC and crystallographic studies in order to ascertain whether any stabilising effect is Y220C specific or mutant independent owing to the high reactivity of a number of p53's cysteine residues.

2.1.3 σ-Hole bonding

The introduction of halogens in drug molecules is widespread with estimations that one or more halogens are contained in around 20% of drugs. Historically, halogenation, in particular fluorination, of compounds in hit-to-lead or lead-to-drug optimisations has sought to modulate the LogP, metabolic stability, basicity and bioavailability of compounds. The so termed 'halogen bonding' or σ-hole bonding phenomenon has attracted growing attention owing to several systematic studies of halogen bonds in protein-ligand complexes deducing that they can serve as a powerful tool with which to increase binding affinity and selectivity. Hence, there has been a steady rise of the use of halogen bonding in rational design of compounds seeking to directly exploit these interactions to improve ligand binding affinity.

The term halogen bonding refers to the intermolecular interaction between a covalently bound halogen atom (RX) and a negatively polarised site (B), such as a lone pair of a Lewis base or an anion (Figure 2.6). It is not a new phenomenon; complexes of Cl_2, Br_2 and I_2 with amines have been observed as early as the 19th century.[292,293] However, particularly important was the observation made by Murray-Rust *et al.,* that, from crystallographic surveys, it appeared that covalently-bound halogens can interact attractively and with high anisotropy with both nucleophiles (B) and electrophiles (E) (Figure 2.6).[294,295] Observations based on computed electrostatic potentials revealed that regions of positive electrostatic potential, σ-holes, along the extensions of the covalent bond are responsible for interactions with donors of electron density (anion or neutral

species possessing at least one nucleophilic region) and areas of negative electrostatic potential that occur equatorially are responsible for interactions with electrophiles (Figure 2.6, B).[296] The term halogen bonding only refers to the former, XB or σ-hole bonding.

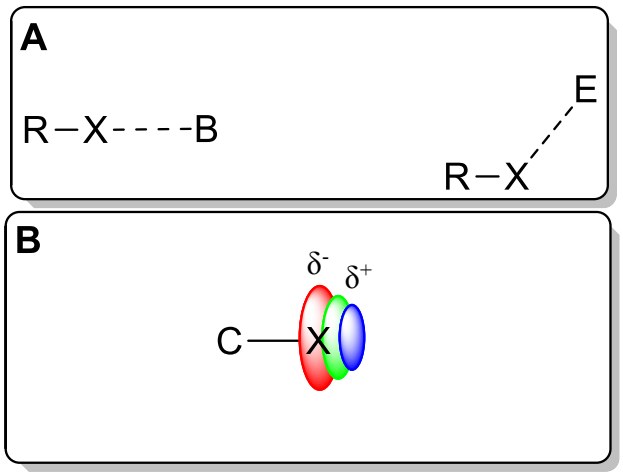


Figure 2.6: Generic depiction of halogen bonding. (A) Representation of how halogens can interact with high anisotropy with negative centers (B) and postitive centers (E). (B) Depiction of electrostatic potential showing the sigma hole. Blue represents electropositive, re represents electronegative. R = any group and X = I, Br, Cl.

The ability of chlorine, bromine and iodine to form halogen bonds is well documented; however, halogen bonding involving fluorine is seldom reported owing to its high electronegativity and lack of polarizability. Recent studies have suggested that under certain circumstances, as a result of the polarization of fluorine's electron cloud by a powerful electron withdrawing group, fluorine has the capability to form halogen bonds.[297–299] Other factors that influence the strength of the interaction include the size of the σ-hole, the angle of the interaction, donor atom properties and the internuclear distance between the halogen and donor. The size of the σ-hole depends on two factors, the polarizability of the halogen and the chemical environment around the halogen. Heavier atoms tend to possess greater polarizability by virtue of their size and hence follows the trend I > Br > Cl > F.[300–302] Additionally, larger σ-holes occur when there is an electron-withdrawing group near the halogen.[303,304]

The angle Θ_1 greatly influences the strength of interaction with the ideal geometry being a 'head-on' contact between the σ-hole of a halogen to the halogen bond acceptor with a R-X---B angle of 180 ° (Figure 2.7). Angle deviations of 25-30 ° correspond to a 50% reduction in halogen bond strength and angle deviations of greater than 40 ° cause a deleterious effect.[305–307]

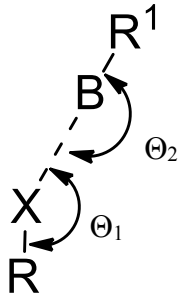


Figure 2.7: Preferred geometry of halogen bonding.

The preferred value for angle Θ_2 depends on the nature of the acceptor atom B, for example, Θ_2 for a carbonyl oxygen would be around 120 °.[307]

2.1.3.1 Sulfur ··· Halogen interactions

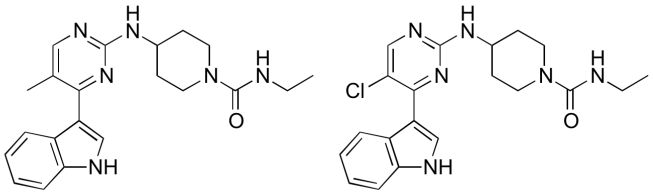
Recent statistical evaluations, in which the PDB has been searched for ligand-protein halogen bonds, have indicated that the two major occurrences of halogen bonding are between the carbonyl of the protein backbone (53%) and X···π interactions (33%). Contacts involving sulfur are far less common (5%).[308] Owing to the serendipitous nature in which most ligand-protein halogen bonds are discovered, and the much higher natural occurrence of carbonyls within protein structures than sulfur containing amino-acids, the relatively low incidence of sulfur-halogen bonding does not necessarily correspond with the interaction being unfavourable.

Quantum chemical calculations using ligand model systems and different molecular representations of methionine using DFT-D, MP2 and CCSD(T) with large basis sets characterized R-X···S interaction strengths and found an order of strength Cl < Br < I.[309] Interestingly, all halogen bonds showed an X···B distance below the Van der Waals radii of both atoms. The equilibrium distances of all X···S bonds were quite similar (3.4 Å) indicating that substitution of Cl for Br or I should be feasible despite the differences in size. Comparison of XB interaction with that of moderate hydrogen bonds, such as a phenol OH···S, considering desolvation, revealed that X···S bonds can outperform hydrogen bonds with complex formation energies of -1.4 kJ/mol and -11.3 kJ/mol respectively.

Although much work has been done computationally to attempt to quantify the strength of S···X, very few examples of S···X bonding have been noted experimentally to increase

ligand-protein affinity. However, the limited SAR associated with **2.4**, an inhibitor of c-jun N-terminal kinases (JNKs) 1-3, suggests a positive effect for S···Cl bonding (Figure 2.8).[310] The interaction formed between the pyrimidine chlorine and the sulfur of the gatekeeper residue Met146 results in IC_{50} of 57 nM vs JNK3 whereas replacing the chlorine with a methyl group results in a 7-fold reduction in potency, with an IC_{50} of 410 nM.

A



2.3 **2.4**

	IC_{50} (nM)	
	2.3	**2.4**
JNK1	320	13
JNK2	250	25
JNK3	410	57

B

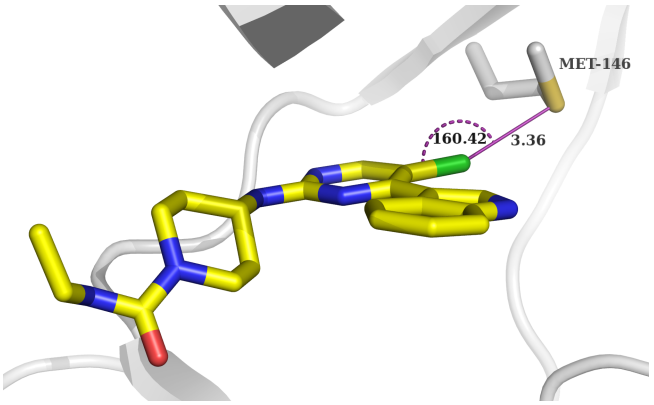


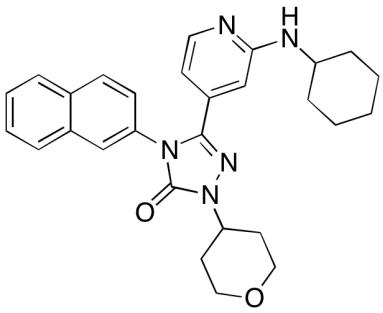
Figure 2.8: Structure of JNK3 inhibitors **2.3** and **2.4**. (A) SAR of two JKN3 inhibitors, (B) Inhibitor **2.4** bound to JNK3 (PDB: 2P33). For clarity, only the gatekeeper residue Met146 is shown. Halogen bond shown in purple with angle C-Cl···S highlighted.

Targeting the Cys220 sub pocket in p53 through halogen bonding, to seek to improve ligand-protein affinity, is particularly attractive owing to the hydrophobicity of the mutation induced pocket. It was previously noted that particularly polar fragments failed to reach Cys220 and hence, groups at R and R^1, which have bond vectors that can target Cys220, should be relatively hydrophobic. As halogenation of compounds is commonly used to modulate LogP, this may present an ideal method of targeting the mutation-induced sub pocket. In addition to targeting cysteine for halogen bonding, backbone carbonyls of residues L145, W146, V147 may be targeted.

2.1.3.2 Sulfur ··· Aromatic interactions

Additionally to halogens, elements of groups IV-VI also possess σ-holes and hence many non-bonding interactions between these elements and donor atoms, e.g. S···O, S···N and S···π can be attributed to σ-hole bonding.[311,312] Sulfur···π bonding has been the topic of many computational and experimental studies by the use of model systems of cysteine (H_2S) or methionine (Me_2S) with aromatic rings.[313–317] Generally, the overall consensus points to a preferred interaction distance of around 5 Å with a preferred geometry consisting of sulfur above the ring plane with either hydrogen atoms or methyl groups more proximal to the π-face and the sulfur atom more distal. These experiments suggest that the S···π interaction is relatively weak with stabilisation energies of 1-2 kcal/mol. Experimental studies, in which S···π interaction energies of methionine or cysteine residues in peptides were assessed, imply a role for other atoms bound to sulfur in the overall free binding energies, for example S-H···π H-bonding.[318,319] As for S···X bonding, examples in the pdb of S···π interactions which contribute favourably to binding affinity of ligand-protein complexes are discovered serendipitously. For example, a specific S···π interaction was highlighted in the discovery of selective JNK2 and 3 inhibitors.[320] A crystal structure of **2.5** bound to JNK3 revealed an S···π interaction between Met146 and the naphthalene moiety, which the authors reasoned was responsible for isoform selectivity as well as potency (Figure 2.9).

A



2.5

	IC_{50} (nM)
JNK1	420
JNK2	97
JNK3	16

B

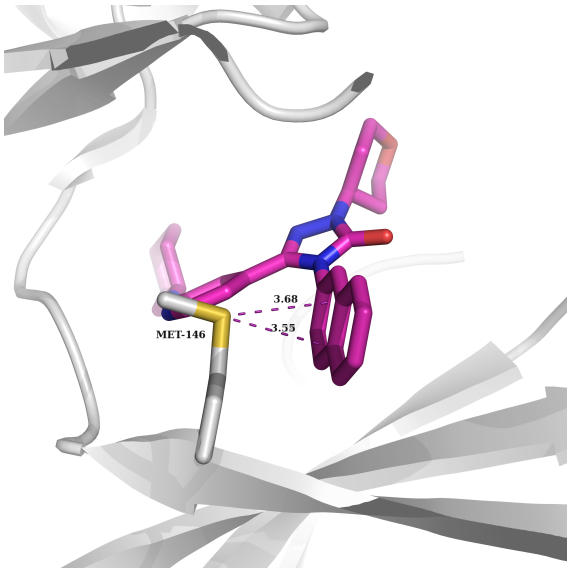


Figure 2.9: Example of S···π interactions of 2.5 with JNK3. (A) SAR of JNK3 inhibitor **2.5**, (B) Inhibitor **2.5** bound to JNK3 (PDB: 3OY1). S···π interactions shown in purple dashed bonds. For clarity, Met146 side chain is shown. Distance of interaction shown on top of each interaction (3.6 and 3.5 Å).

By incorporation of aromatic groups at R, a complementary strategy to S···X bonding may be considered in PK083 analogue design. Further rationalization of this approach from existing SAR in the mutational cavity is offered in Section 2.3.

2.2. Multipolar Interactions

As mentioned in Section 2.1.3, fluorine rarely forms halogen bonds, however there are numerous examples of direct protein-fluorine interactions reported.[321,322] A scan of the PDB reveals that fluorinated ligands are frequently found to interact with proteins in different ways. This may be through polar interactions with hydrogen bond donors, hydrophobic interactions with lipophilic side-chains or orthogonal multipolar interactions.[323] Orthogonal multipolar interactions are typified by a close orthogonal contact between two dipolar functional groups, for example, C=O and C-F.[324] Theoretical studies have deduced that, in an orthogonal relationship, the dipole contribution to interaction energy is zero such that higher order electrostatic and dispersion terms must be responsible for the attractiveness of the interaction.[325] Fluorine and other halogens can interact with carbonyls via the carbon or oxygen via multipolar interactions or halogen bonding respectively depending on the geometry of the interaction and the halogen involved. Surveys of the Cambridge structural database (CSD) for C=O···X where X-C or X-O is less than the Van der Waals contact reveals information about the nature of these interactions.[326] For fluorine, there is no orientation dependence and appears to approach the carbonyl from any angle. However, for chlorine there are two distinct distributions for halogen bonding and multipolar interactions which appear to occur at roughly the same frequency. Halogen bonding, as discussed previously, has a preferred angle C-X···B of between 160-180° corresponding to a roughly in-plane orientation. Multipolar interactions have a preferred angle C=O···X of between 60-80°, adopting a side-on above the plane orientation (Figure 2.10).

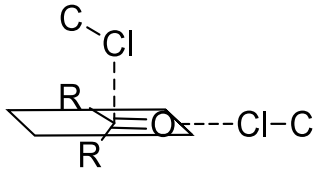


Figure 2.10: Difference in geometry between halogen bonding and multipolar interactions formed by chlorine.

There are numerous examples of SAR's which link multipolar interactions with positive effects on free binding energy. The Abelson murine leukemia viral homolog 1 (abl) kinase inhibitor Nilotinib forms a multipolar interaction between the CF_3 moiety the backone

C=O of Ala380 as well as polar interactions with His361 N-H and the C-H of Ile293 (Figure 2.11).[327] The CF_3 is 5-fold more active than its methylated analogue in an autophosphorylation assay.

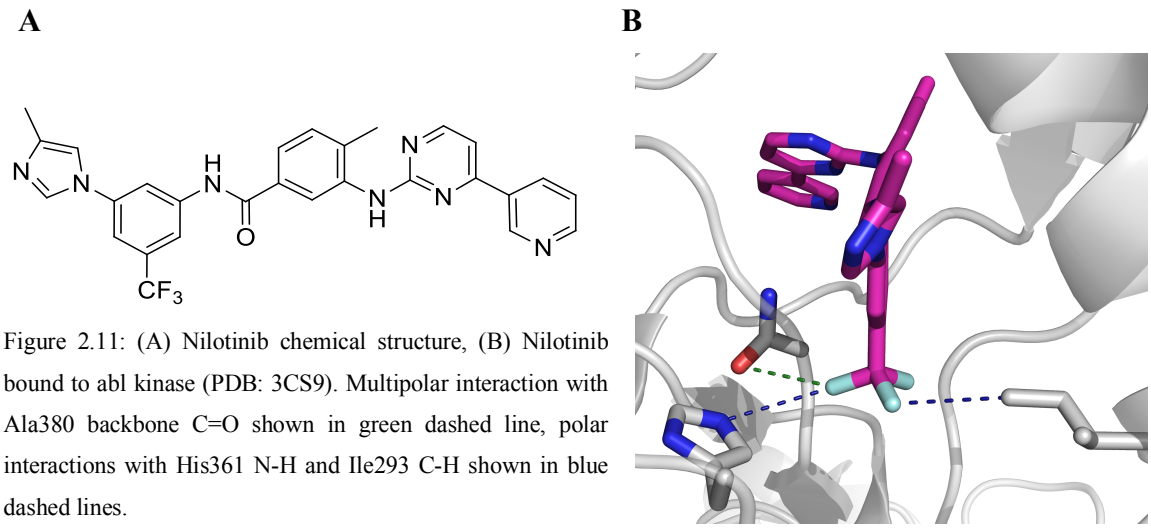
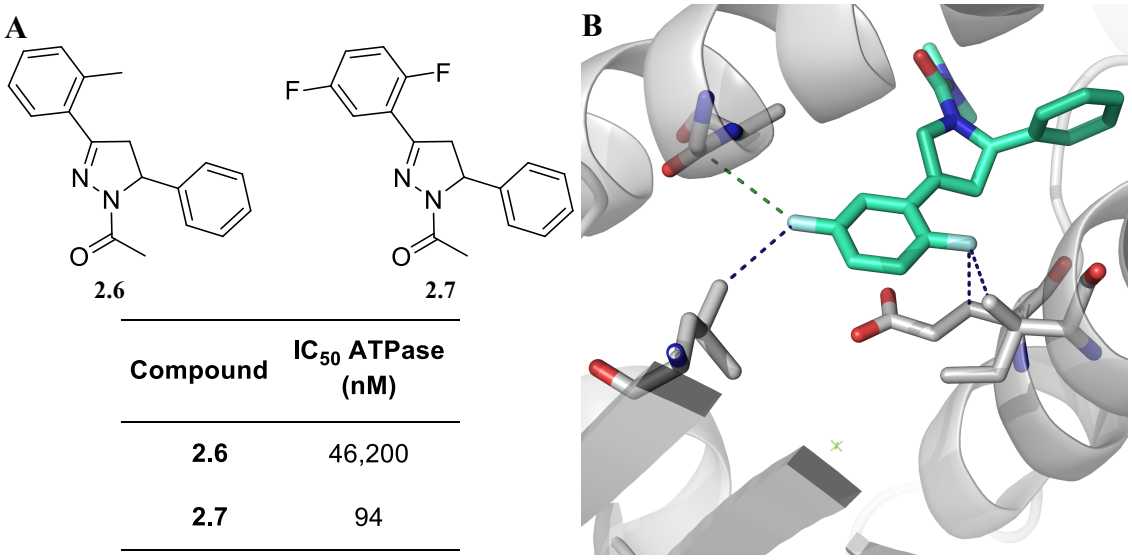


Figure 2.11: (A) Nilotinib chemical structure, (B) Nilotinib bound to abl kinase (PDB: 3CS9). Multipolar interaction with Ala380 backbone C=O shown in green dashed line, polar interactions with His361 N-H and Ile293 C-H shown in blue dashed lines.

Inhibitors of kinesin spindle protein (KSP) display a dramatic increase in affinity for fluorinated analogues.[328] Figure 2.12 shows the increase in potency between the methylated (**2.6**) and fluorinated forms (**2.7**). The aryl-fluoro ring forms a multipolar contact with Gly217.



Compound	IC_{50} ATPase (nM)
2.6	46,200
2.7	94

Figure 2.12: (A) SAR of KSP inhibitors **2.6** and **2.7**. (B) X-ray of **2.7** bound to able kinase (PDB: 2FL6). Multipolar interaction for F with C=O of Gly217 shown in green dashed line. Polar interactions with Glu116, Ile136 and Leu160 shown in blue dashed lines.

It is clear that orthogonal multipolar interactions can positively impact free binding energies however, it remains to be ascertained whether this is solely due to multipolar contacts or is more complex involving changes in residual mobility, desolvation and

hydrophobic interactions. Perhaps, rather than seeking to target the protein backbone for multipolar interactions, potentially 'fluorophillic' pockets should be identified and targeted according to a set of criteria involving hydrophobics, solvation among other parameters.

2.3 Existing SAR in the Central Cavity - Fragments

Basse *et al*., explored the ligand binding sites of T-p53-Y220C, an engineered stable variant of p53 possessing the Y220C mutation, by screening against a chemically diverse fragment library of 1900 compounds using waterLOGSY and thermal denaturation scanning fluorimetry.[291] They identified 252 hits, of which 87 were confirmed using $^{15}N/^{1}H$ HSQC (Table 2.3). By examination of structural similarities, such as: the number of rings in the central scaffold and the spacing between them, the aromaticity of the ring system, the number of molecules in each ring, and the size of substituent groups, they could be grouped into 7 unique groups depending on their displayed pharmacophore (Figure 2.13).

Variable	WaterLOGSY	Thermal Denaturation Scanning Fluorimetry
Number of fragments screened	1895	1895
Number of identified hits	205	47
Number of hits confirmed by $^{15}N/^{1}H$ HSQC	70	17
Final hit ratio, %	3.7	0.9

Table 2.3: Fragment hits from WaterLOGSY and thermal denaturation scanning fluorimetry.[291]

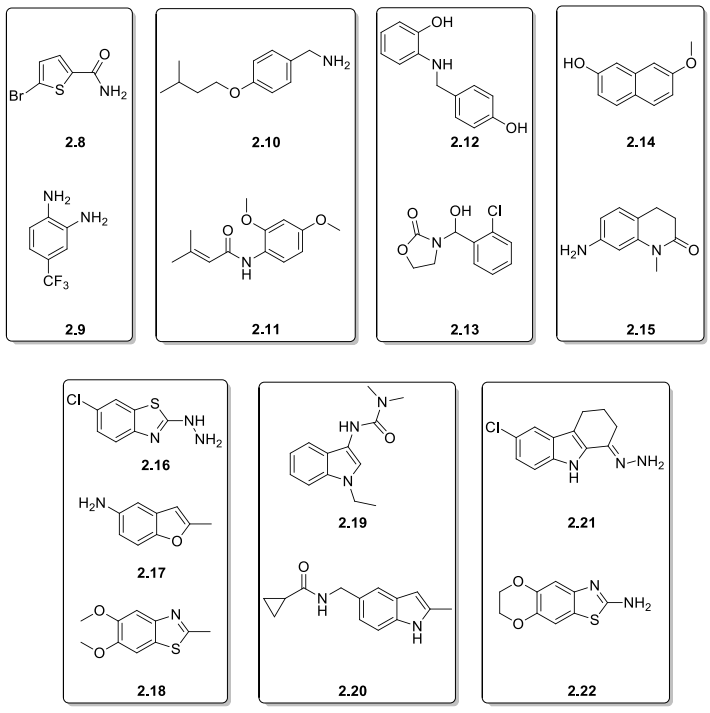


Figure 2.13: Fragments hits grouped according to identified pharmacophore.

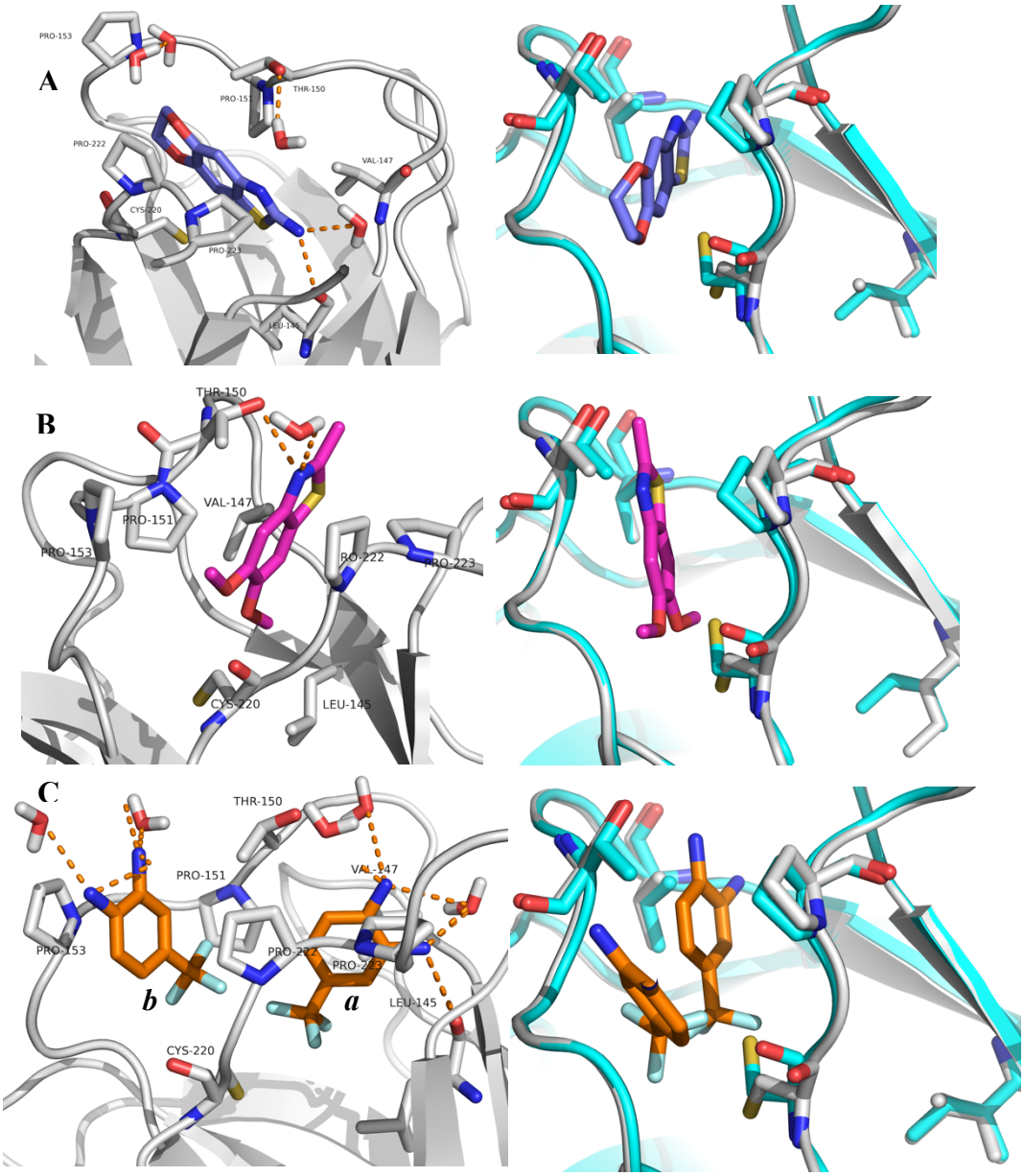


Figure 2.14: Binding modes of fragments **2.22**, **2.18** and **2.9**.(A) Fragment **2.22** bound to the Y220C pocket (PDB: 2X0U) and the comparison of the residues in the bound protein (grey) and ligand-free protein (PDB: 2J1X, cyan); (B) Fragment **2.18** bound to Y220C (PDB: 2X0W)and comparison of Y220C in the bound protein (grey with ligand-free protein (cyan); (C) Fragment **2.9** bound to the Y220C pocket (PDB: 2X0V) and comparison of key residues in the bound protein (grey) and ligand-free protein (cyan). Hydrogen bonds are shown as orange dashed lines.

For fragments **2.9**, **2.18** and **2.22** co-crystal structures were solved which identified their mode of action (Figure 2.14). Generally, the fragments' aromatic rings are sandwiched between Pro151 and Val147 or Pro222 and Pro223. In order to accommodate the ligands, the cysteine flips into a different conformation to that which is observed in the native, ligand-free structure, deepening the pocket by a maximum of 2 Å. This cysteine flip also causes Val147, Ile 232, Thr150 and Pro222 to shift. Interestingly, for fragment **2.9**, the

pocket is occupied by two ligand molecules at opposite ends which both have the CF_3 group oriented towards Cys220. Ligand *a* sits between Val147 and Pro222 forming hydrogen bonds between the aniline and the C=O of Leu145 and a conserved water and the other, *b*, sits between Pro153 and Pro222. The CF_3 group orientation towards the cysteine hints at the mutation induced sub-pocket as possessing a 'fluorophillic' characteristic.

This fragment screen indicates a role for exploiting p53-Y220C's structural plasticity by designing molecules which target the transiently open subsite III. For example the fragments **1.1** and **1.2** do not induce cysteine flipping, however, their structural analogues **1.3** and **1.4** both exhibit induced-fit ligand binding (Figure 2.15).[258]

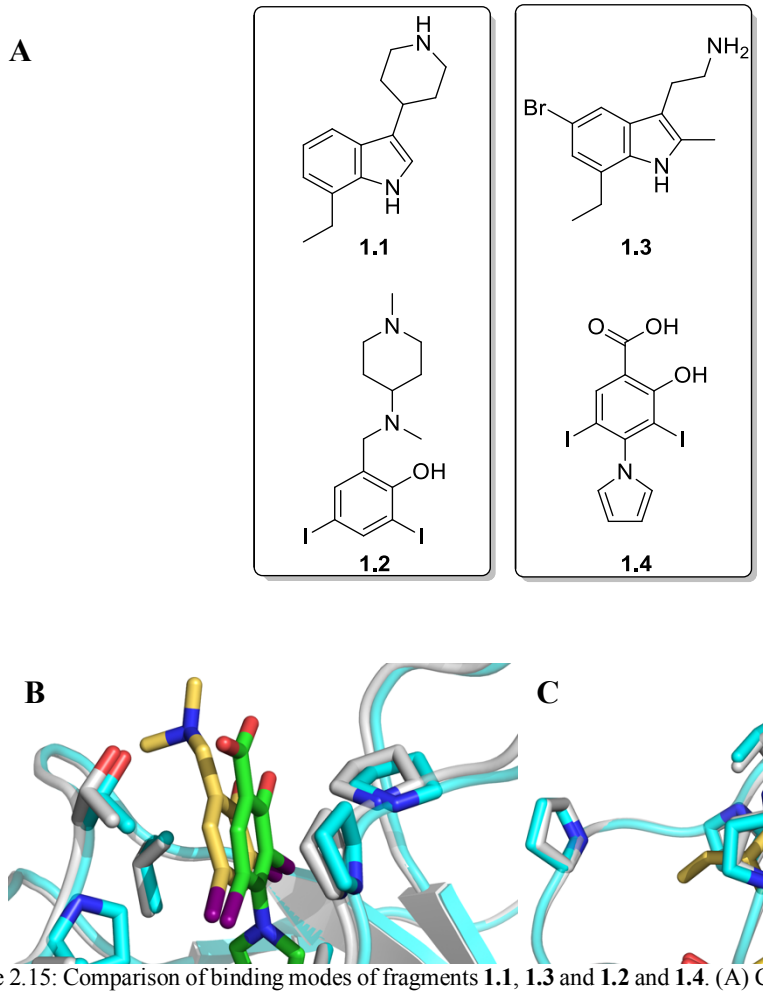


Figure 2.15: Comparison of binding modes of fragments **1.1**, **1.3** and **1.2** and **1.4**. (A) Chemical structure of fragments **1.1**, **1.2**, **1.3** and **1.4**. (B) Comparison of binding modes of fragments **1.2** (PDB: 4AGL) and **1.4** (PDB: 5AOJ) in p53-Y220C. Colour scheme: **1.2**, yellow (carbon), grey (p53-Y220C), **1.4**, green (carbon), cyan (p53-Y220C). (C) Comparison of binding modes of fragments **1.1** (PDB: 5AB9) and **1.3** (PDB: 5AOI) in p53-Y220C. Colour scheme: **1.1**, orange (carbon), grey (p53-Y220C), **1.3**, yellow (carbon), cyan (p53-Y220C). Note the cysteine position is altered (comparison of grey vs cyan) in both B and C.

Fragment **1.4** adopts the same binding pose as **1.2**, where the aromatic ring is sandwiched between Pro22, Pro223, Pro151 and Val147. The pyrrole substitution ortho- to both iodine atoms in fragment **1.4** forces the mutant specific Cys220 into the 'open' conformation. In this example, this is a favourable effect as there is an observed increase the K_D from 820 μM in fragment **1.2** to 21 μM for fragment **1.4**. Conversely, the SAR for indoles **1.1** and **1.3**, the effect of introduction of bulky hydrophobic groups, bromine on the benzene ring and methyl on the indole, results in a shift in binding position. Fragment **1.1** sits between Pro223 and Val147 where the ethyl group is buried in the deep hydrophobic mutation site. Fragment **1.3** flips position with the ethyl group pointing towards subsite II and the bromine atom oriented toward Cys220, forcing it into the 'open' conformation. The overall effect is a reduction in affinity with K_D's of 470 μM and 940 μM for fragments **1.1** and **1.3** respectively.

2.4 Design of PK083 Analogues

Comparison of PK083 and fragment **2.22** in the Y220C pocket reveals that the central aromatic core of PK083 is shifted towards subsite I compared with fragment **2.22** (Figure 2.16). The carbazole aromatic system sits much higher in the pocket compared to the benzothiazole scaffold of fragment **2.22** owing to the functionalisation of the central nitrogen. The effect of this is that the rigid benzothiazole scaffold forces the cysteine to flip into the 'open' conformation whereas the cysteine in the PK083-Y220C structure sits in the 'closed' conformation, albeit with some fluxionality.

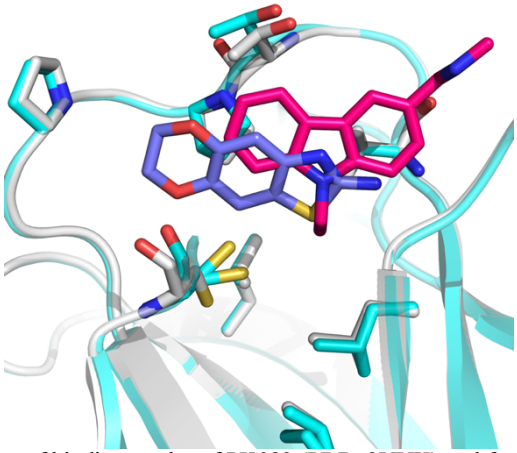


Figure 2.16: Comparison of binding modes of PK083 (PDB: 2VUK) and fragment **2.22** (PDB: 2XOU) in the central cavity and subsite III.

The fluxionality of the cysteine in PK083-Y220C suggests that there is potential for growth at R through the introduction of hydrophobic groups. Hence, the first method of targeting the mutational cavity should be to establish, through SAR, what is the maximum size hydrophobic group at the central nitrogen tolerated?

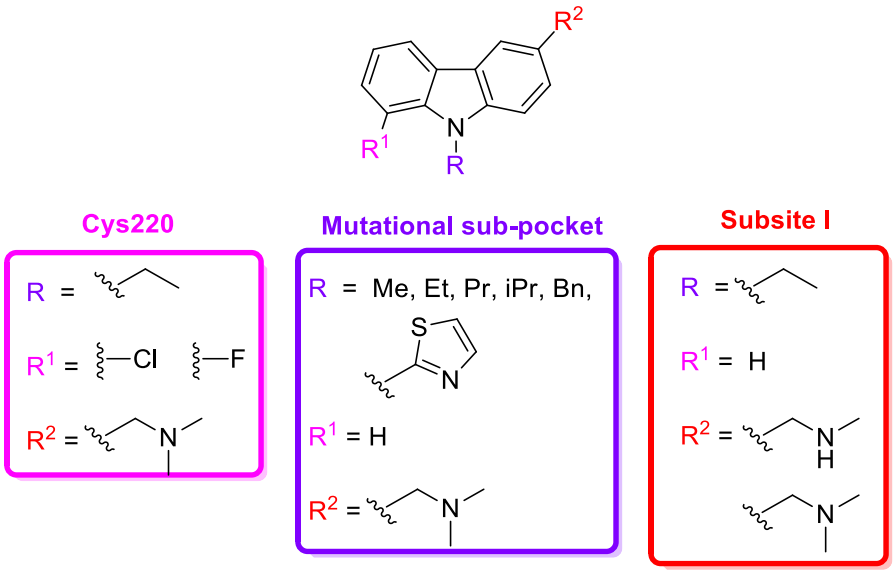


Figure 2.17: Summary of compounds synthesised.

To address this question, two approaches were adapted (Figure 2.17). Firstly, through sequential chain growth at R by introduction of larger flexible aliphatic hydrocarbon groups and secondly, through the introduction of rigid aromatic groups. The latter aimed at affinity increases through targeting the transiently open subsite III, as well as seeking for potential S···π interactions.

From the identification of the Cys220 sub-pocket as 'fluorophillic' through the binding of fragment **2.9**, fluorinated analogues of PK083 were designed seeking to increase affinity through fluorination of R^1. Additionally, Cys220 was targeted directly for S···X bonding through chlorination at R^1 (Figure 2.17).

Furthermore, by variation of the benzylic amine at the solvent exposed subsite I (R^2), additional SAR could be established which can inform the effect of the amino-exposed side-chain on overall potency (Figure 2.17).

Covalent modification compounds were designed and synthesised by Dr's A. Close and B. Springett and are not discussed in this work.

2.5 Carbazole – Discovery, Applications and Chemistry

Carbazole was first described in 1872 by Graebe and Glaser who obtained significant quantities from the anthracene fraction of coal tar distillate.[329] However, it was not until the antimicrobial properties of murrayanine (3-formyl-1-methoxycarbazole), which was isolated from the plant *Murraya koenigii*, were discovered that there was interest from chemists and biologists in this privileged structure.[330–332] Since then, naturally occurring carbazoles and synthetic derivatives have formed the scaffold of approved drugs and many 'lead' compounds targeting a plethora of diseases (Figure 2.18). For example, the marketed drug carprofen is a non-steroidal anti-inflammatory drug used as a supportive

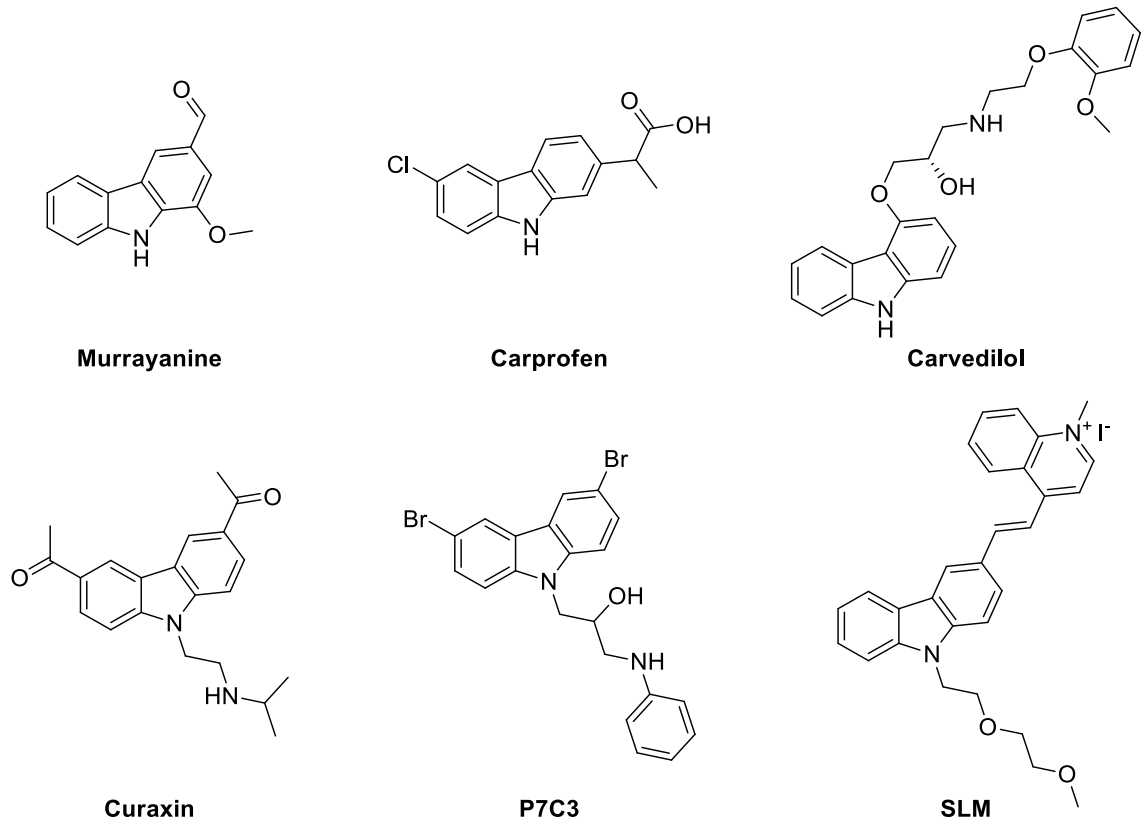


Figure 2.18: Chemical structures of pharmacologically active carbazoles.

treatment for various conditions in animals.[333] Carprofen reduces inflammation by inhibition of COX-1 and COX-2. The Curaxins are a family of carbazole compounds being developed as an anti-cancer therapy that bind the heterodimeric protein complex facilitates chromatin transcription (FACT), the sequestration of which causes p53 activation and NF-κB inhibition leading to tumour cell death without DNA damage. [334] Carvedilol (Coreg) is a non-selective β- and α-adrenergic blocker used to treat cardiovascular disease.[335] Carbazole derivatives have also been discovered as potential treatments of neurodegenerative diseases such as Parkinson's disease. The small

molecule P7C3 (an aminopropyl carbazole) blocks 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) mediated cell death of dopaminergic neurons in the substantia migra of adult mice, a model for Parkinson's disease.[336] Additionally, a carbazole based fluorophore SLM is being investigated as a potential treatment for Alzheimer's disease.[337] Aggregation of amyloid β to form neurotoxic plaques leading to cognitive deficits is one possible mechanism of Alzheimer's disease. SLM binds and inhibits aggregation of amyloid β. Owing to the pharmacological importance of carbazole synthetic derivatives and natural products, there has been an enormous development in chemistry towards the synthesis of highly functionalized carbazole scaffolds.

2.5.1 Functionalization of Carbazole

Owing to its highly electron rich nature, carbazole is a modest nucleophile that can be readily functionalized with a wide variety of electrophiles. Its reactivity can be viewed analogously to diphenylanilines, in which the central nitrogen group directs *ortho*- or *para*- the central nitrogen.[338] The electrophilic aromatic substitution of carbazole leads to 1,3,6,8 substituted scaffolds (Figure 2.19). Functionalization of the 9- position can be achieved through treatment with base to afford a nitrogen nucleophile which reacts with a plethora of electrophiles to generate N-alkylated carbazoles. The nuances in carbazole reactivity towards electrophiles are discussed below.

2.5.1.1 Electrophilic Aromatic Substitution

Carbazole is reactive towards a plethora of electrophilic aromatic substitution reactions, for example, bromination, nitration, etc. The most reactive positions are 3 and 6 followed by 1 and 8. Owing to the high level of symmetry, formation of poly-substituted carbazoles is common depending reaction conditions. For example, bromination of carbazole with N-bromosuccinimide (NBS) can be controlled by varying the molar equivalents of NBS (Scheme 2.2).[339]

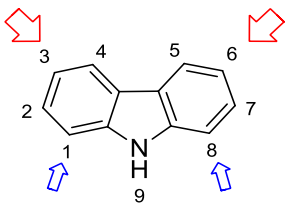
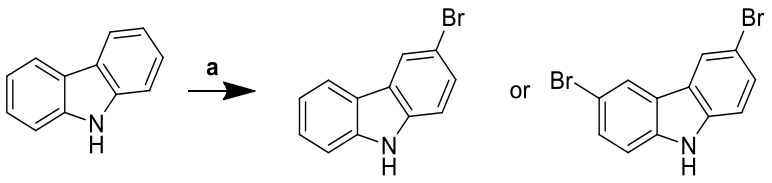
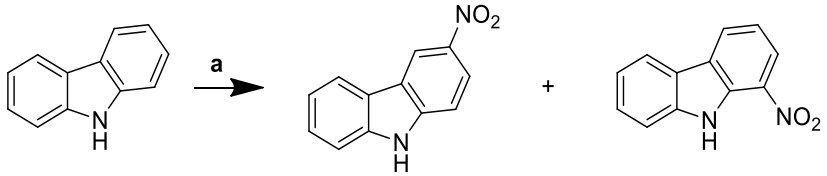


Figure 2.19: Reactivity of the carbazole ring towards electrophilic aromatic substitution reactions.



Scheme 2.2: Bromination of carbazole. (a) NBS, DCM, DMF, rt. One equivalent of NBS gives 95% 3-bromocarbazole, two equivalents of NBS gives 96% 3,6-dibromocarbazole

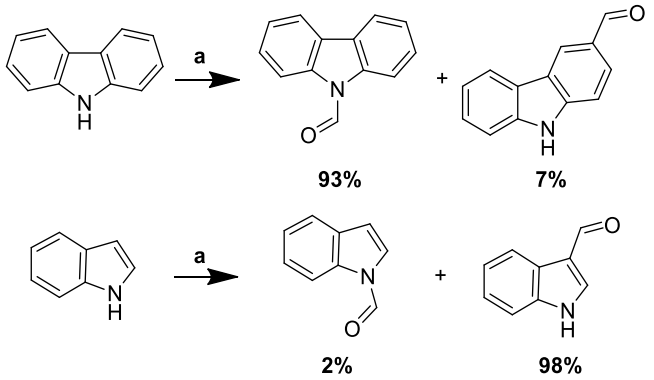
However, when the reaction conditions are severe, as in nitration, control over regio-selectivity is challenging and the formation of a mixture of isomers is common (Scheme 2.3).[340]



Scheme 2.3: Nitration of carbazole. (a) HNO_3, AcOH, 50 °C, 3 h.

The reaction is suggested to go through formation of a 9-nitroso intermediate, which, under the strong oxidizing conditions is converted to a 9-nitro intermediate which rapidly rearranges to form 1- or 3-nitrocarbazole in a 3:7 ratio. By alkylation of position 9, the reaction proceeds with greater regio-isomeric control (1:9) via direct nitration of the aromatic ring in contrast to nitrosation of free carbazole (NH).[340]

Furthermore, examination of indole and carbazole's reactivity towards Vilsmeier-Haack formylation and Mannich aminomethylation reveals a much higher preference for reactivity at the N-H of carbazole over the N-H of indole (Scheme 2.4). Vilsmeier-Haack formylation of both indole and carbazole yields N-acylated products however, the reaction with carbazole yields almost exclusively N-acetylated product, whereas, for indole, N-aceylation only accounts for 2% of the overall yield.[341]

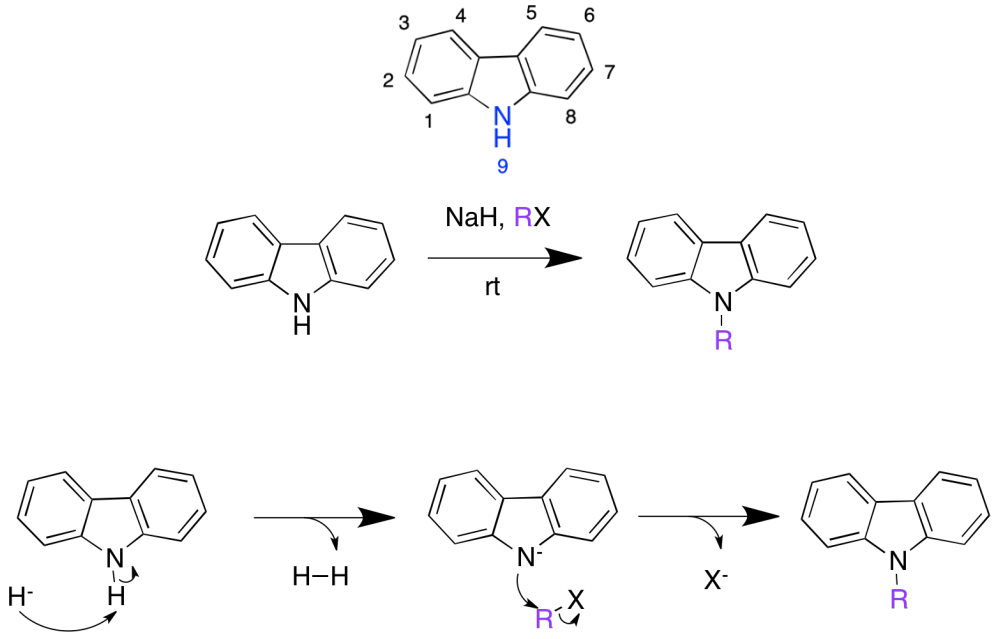


Scheme 2.4: Reactivity of carbazole and indole towards Vilsmeier-Haack acetylation.

This pattern of reactivity is extended to Mannich aminomethylation reactions. Indole reacts with Eschenmoser's salt to yield both N- and C-substituted benzylic amines in a ratio of 1:7.7, whereas, carbazole exclusively reacts to form the N-substituted derivative.[342,343] Kinetics studies have calculated rate constants of N-acteylation of indole (3.3 X 10^4 mol^{-1} s^{-1}) and carbazole (24.3 mol^{-1} s^{-1}) indicating carbazole undergoes N-acetylation ca. 7 times faster than indole.[344] The authors attributed this to the differences in the overall loss of aromaticity in the transition state as approximated by the Wheland intermediates. The nitrogen quaternisation arising from the direct attack of the electrophile on the nitrogen atom leads to a larger degree of bond fixation in indole (only one aromatic ring to delocalize into) than in carbazole (two aromatic rings).

2.5.1.2 Nucleophilic Substitution Reactions of Carbazole

Carbazole is reactive in nucleophilic substitution reactions at position 9. The pKa of carbazole (N-H) is around 20 and hence is weakly acidic and can only be deprotonated (N-H) by strong bases such as sodium hydride. The generated nitrogen nucleophile can then undergo a nucleophilic substitution reaction with an electrophile, such as an alkyl halide to give N-functionalised carbazoles (Scheme 2.5).



Scheme 2.5: Reaction of carbazole nucleophile with electrophile to yield N-functionalized products.

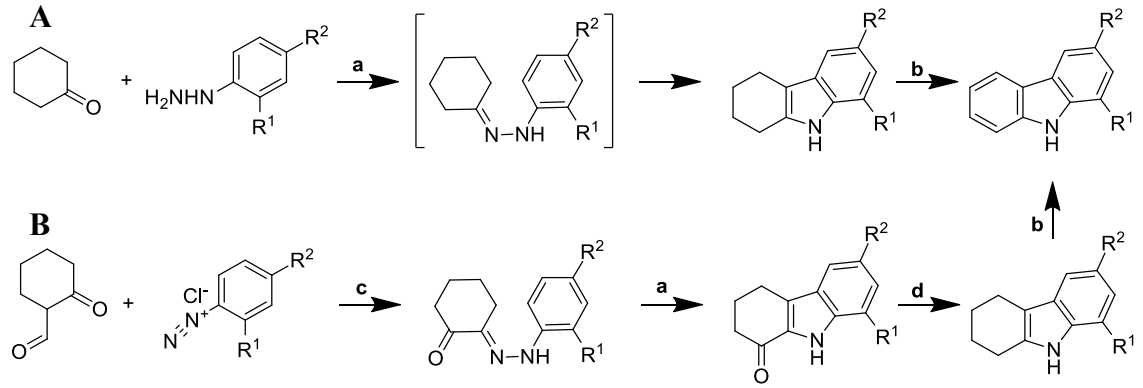
2.5.2 Synthesis of Carbazole

Functionalization of the carbazole scaffold is limited to the formation of 1,8- and 3,6-substituted carbazoles. If functionalization of positions 2,4,5 or 7 of the ring is required or unsymmetrical substitution patterns then synthesis of the carbazole scaffold bearing the desired substitution pattern is required. Numerous methods of carbazole synthesis that exhibit a broad range of different frameworks and functional groups have been developed. By far the most common method of carbazole synthesis is by construction of the central pyrrole ring from biaryls with an ortho- substituted nitrogen substituent or diarylamines. Named reactions are discussed below.

2.5.2.1 Fischer-Borsche Synthesis

A large number of carbazole syntheses involve the formation of 1,2,3,4-tetrahydrocarbazole as an intermediate which is then reduced using chloranil or palladium on carbon to yield carbazole.[345,346]

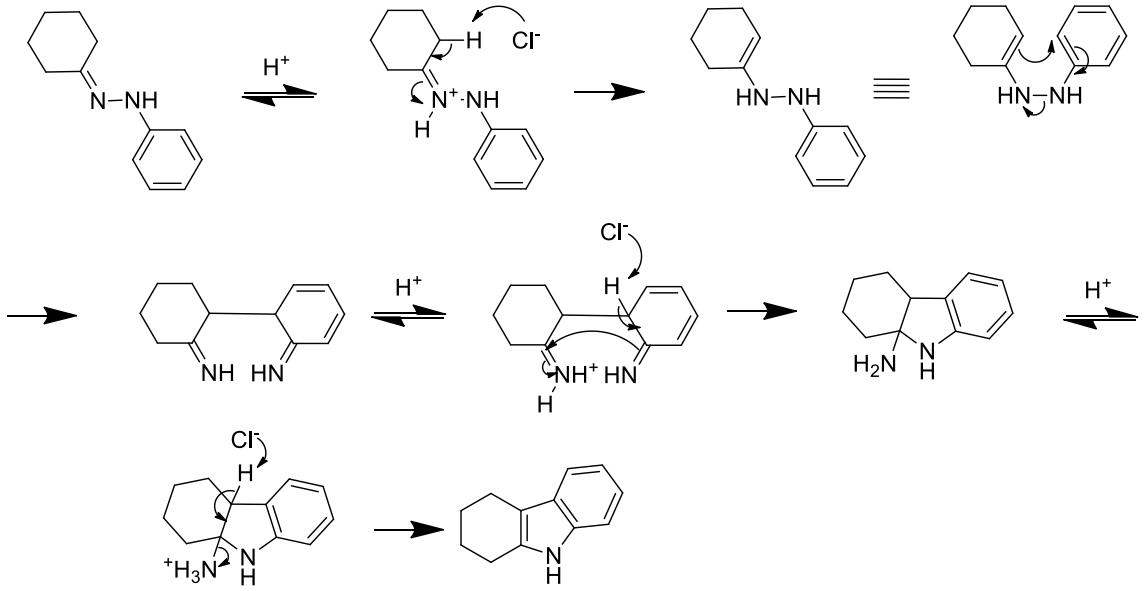
The Fischer method of indole synthesis by indolization of an arylhydrazone by treatment with an acid catalyst was applied to the synthesis of tetrahydrocarbazoles by Borsch (Scheme 2.6A).[347,348] Condensation of cyclohexanone with phenyl hydrazines affords aryl hydrazones which, upon protonation by an acid catalyst, forms a new C-C bond via [3,3]-sigmatropic rearrangement and elimination of ammonia to afford tetrahydrocarbazoles (Scheme 2.7). [349]



Scheme 2.6: (A) Fischer-Borsche synthesis of carbazole from cyclohexanone and phenyl hydrazine. (B) Fisher-Borsch synthesis of carbazole using the Japp-Klingemann reaction to form the aryl hydrazine intermediate. (a) aq AcOH, reflux, (b) chloranil or Pd/C. (c) aq NaOAc, MeOH, (d) $N_2H_4.H_2O$, KOH, reflux.

Alternatively, the hydrazone can be prepared by a Japp-Klingemann reaction by retro-Claisen condensation of 2-formylcyclohexanone with aryldiazonium salts (Scheme 2.6B).[350–353] The resulting hydrazine undergoes Fischer-Borsche cyclization followed by

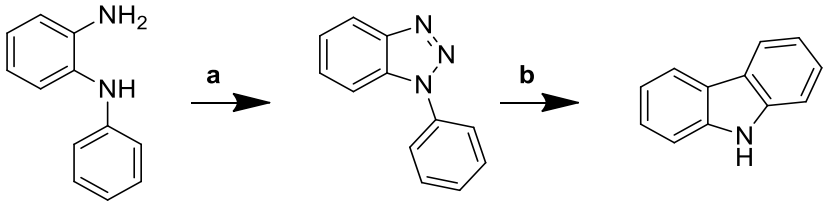
deoxygenation to yield tetrahydrocarbazole which can be reduced to carbazole as previously. The Japp-Klingemann hydrazone synthesis increases the yield of the aryl hydrazone.



Scheme 2.7: Mechanism of the Fischer-Borsche carbazole synthesis.

2.5.2.2 Graebe-Ullmann Synthesis

The Graebe-Ullmann synthesis of carbazole, where N-phenyl-1,2-diaminobenzene reacts with nitrous acid to form 1-phenylbenzotriazole, which is unstable and undergoes thermolysis to yield carbazole, proceeds in almost quantitative yield (Scheme 2.8).[354–356] However, the reaction is very sensitive to substitution and hence is not commonly used to prepare highly substituted carbazole scaffolds. Only a little is known about the mechanism of the reaction, but it most likely a diradical intermediate is involved in the thermolysis of triazole.

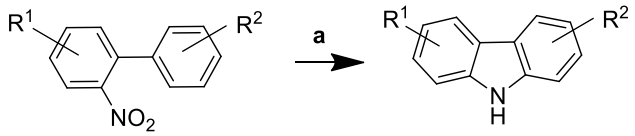


Scheme 2.8: Graebe-Ullmann synthesis of carbazole from N-phenyl-1,2-diaminobenzene. (a) HNO_2; (b) neat, Δ.

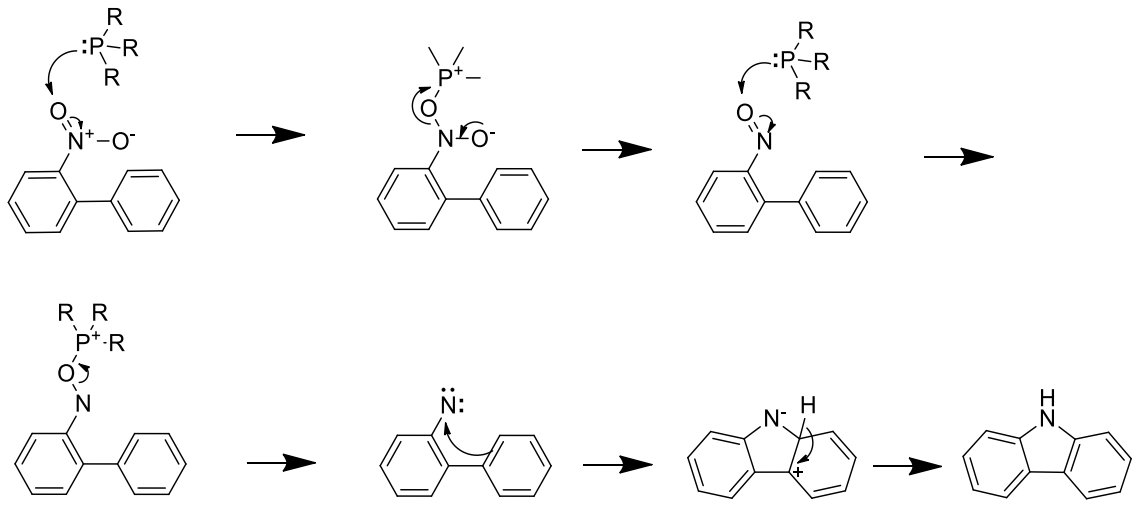
2.5.2.3 Cadogan Synthesis

The reductive cyclization of *ortho*-nitrobiphenyls in triethylphosphite under reflux is known as the Cadogan synthesis (Scheme 2.9).[357] The widely accepted mechanism for this transformation involves the exhaustive deoxygenation of the nitro group to a nitrene than undergoes C- H insertion (Scheme 2.10).[358,359] The reaction has high functional

group tolerance allowing the formation of highly substituted carbazoles. However, the resulting compound is often contaminated with N-ethylated carbazole derivatives which are generated from the reaction of triethyl phosphite or the triethyl phosphate by-product from the reaction and the resulting mixture is often difficult to purify. Recently, this transformation has been carried out more efficiently using carbon monoxide or other organophosphorous reagents such as triphenyl phosphine as stoichiometric reducing agents.[360,361] The most challenging derivatives to couple are those bearing powerful *ortho*- electron-withdrawing groups. However, through microwave-assisted chemistry these compounds can be accessed.



Scheme 2.9: Cadogan synthesis of carbazole starting from an *o*-nitrobiphenyl. (a) $P(OEt)_3$, reflux.



Scheme 2.10: Mechanism of Cadogan reductive cyclization reaction.

2.6 Microwave-assisted organic synthesis

There are very few methods of speeding up an organic reaction; one such way is using microwave-assisted organic synthesis (MAOS). Typically, MAOS is described as high-yielding. This is theorized to occur due to the fact that uniform heating is observed as the reaction vessel is not heated directly (Figure 2.20).[362] Instead, the microwave radiation passes through the walls and heats the solvent and reactants directly. This can lead to less by-products and/or decomposition products.

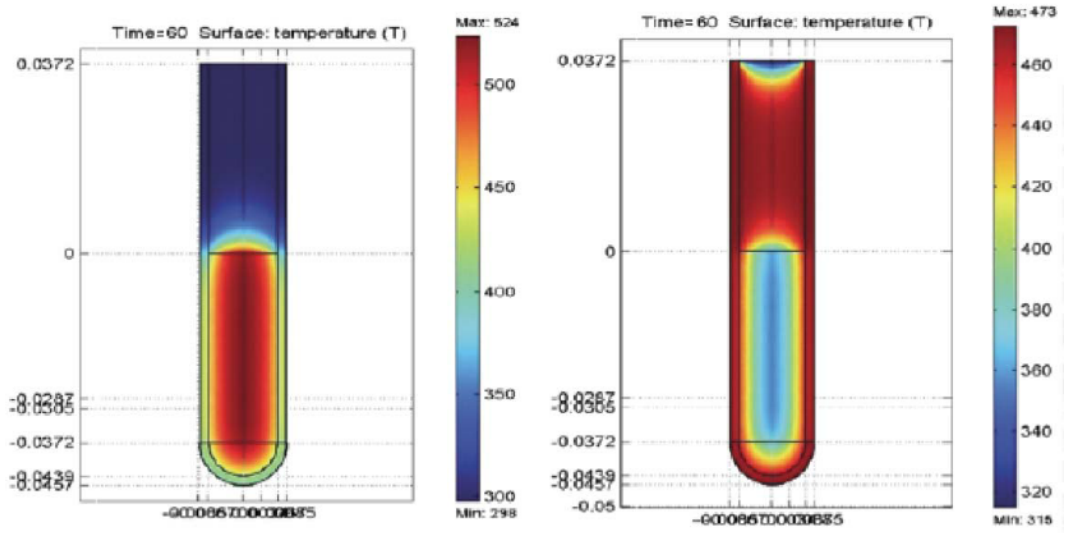


Figure 2.20: The temperature profile after 60 s as affected by microwave irradiation (left) compared to treatment in an oil bath (right).[362] Reproduced from Ref. 362 with permission from the Royal Society of Chemistry.

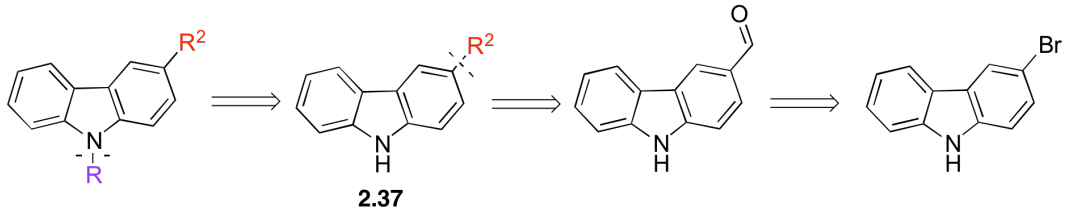
Microwave irradiation is composed of an oscillating electric and magnetic field with wavelengths between 300 MHz and 300 GHz. In an organic reaction, the reaction medium (reactants, solvent, and substrate) is believed to interact with the electric field component. These interactions may be attributed to dipolar polarization and conduction mechanisms.[363] Dipolar polarization generates energy by the mismatch created as a molecular dipole oscillates slightly out of phase with the electric field component. Conduction mechanisms originate from collision theory and describe the increased expenditure of energy related to the rapid movement of ions in a superheated medium. The ability of a solvent to absorb microwave energy and convert this energy into heat can be summarized by a specific loss angle (incorporation of a loss factor and the dielectric constant of the solvent). This is turn is dependent on a specific solvent relaxation time (time it takes for one molecule to return to 36.8% of its original situation when the electric field is turned off). The relaxation time is temperature dependent and decreases as temperature increases. Since the loss angle is dependent on the relaxation time, the ability of the solvent to convert microwave energy into heat will be dependent on the frequency and temperature. This phenomenon is called superheating and results in a solvent being heated to 26°C above its conventional boiling point and is largely believed to be responsible for the rate increases observed.

2.7 Results and Discussion

Firstly, compounds which sought to inform the influence of the benzylic amine were synthesised preferentially to dictate the nature of R^2 for the synthesis of all subsequent compounds. Secondly, the maximum tolerance of the mutational cavity was sought.†

2.7.1 Retrosynthetic Disconnection of Target Compounds

Disconnection of generic PK083 gives two simple routes to compounds which show diversity at R and R^2 in Scheme 2.11. The synthesis starts from the commercially available 3-bromocarbazole. Lithium-halogen exchange followed by reaction with DMF is reported to give 9*H*-carbazole-3-carbaldehyde in moderate yields (50%).[364] Reductive amination with the appropriate amine gives intermediate **2.37** which, following treatment with base, undergoes nucleophilic substitution reaction with a range of electrophiles to introduce diversity and give PK083 analogues. Alternatively, for R = heteroaromatic, copper-mediated cross coupling reactions have been reported for carbazole between the generated nitrogen nucleophile and bromo- or iodoinated heterocycles such as thiophene.[365]



Scheme 2.11: Retrosynthetic disconnection towards compounds targeting subsite I and the mutational sub-pocket.

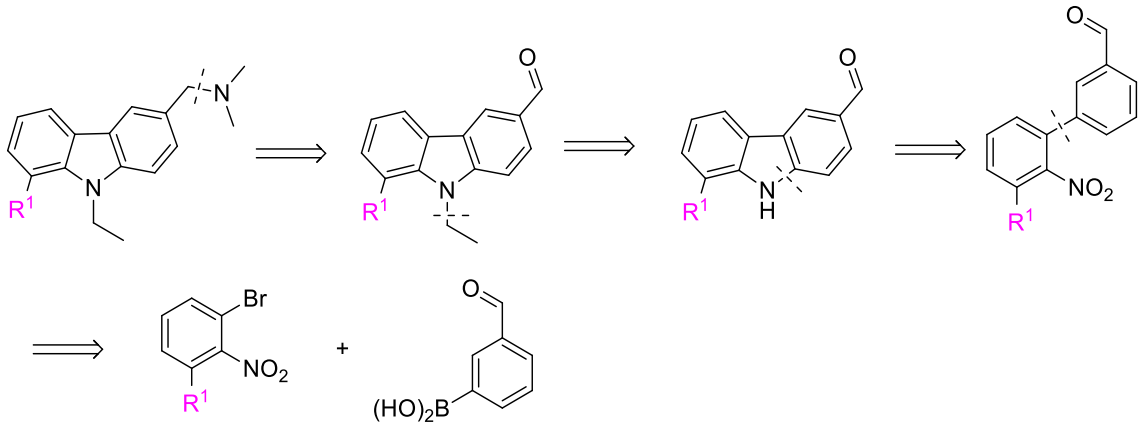
When R = Me, Et, Pr, iPr, Bn, the N-functionalized 3-aldehyde intermediate was commercially available and hence the synthesis was simplified to a one-step reductive amination towards target compounds.

Disconnection of Cys220 targeting compounds (Figure 2.17, pink box) with a 1,6,9-substitution pattern results in the 4-step synthesis described in Scheme 2.12. Suzuki

† Contributions to collaborative research summarized in section 2.7 is as follows. Compound MB016 was synthesised by Matthias Baud (LMB) and PK9277 was synthesised by Bradley Springett (UoS). All other compounds were designed and synthesised by myself. DSF and HSQC measurements on compounds with the identifier PK80xx were carried out by Rainer Wilcken. All other DSF and ITC measurements were carried out by Matthias Bauer (LMB).

coupling of 3-formylphenyl boronic acid with fluoro- or chloro- aromatic nitro bromides yields *o*-nitrobiaryl intermediates.[366]

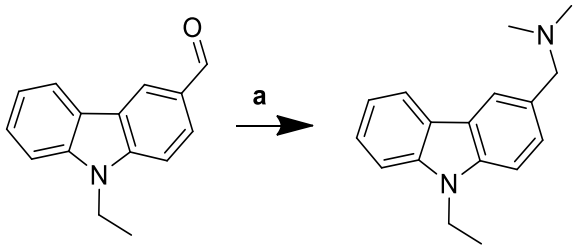
Cadogan cyclization yields the unsymmetrically substituted carbazoles which are then N-ethylated using a nucleophilic substitution reaction to yield 1,6,9-substituted carbazole, that, upon reductive amination yields the target compounds.[361,367]



Scheme 2.12: Reterosynthetic disconnection of Cys220 targeting compounds described in Figure 2.17, pink box.

2.7.2 Synthesis and SAR of a Subsite I Target Compound

Starting from 9-ethyl-carbazole-3-carbaldehyde, which was obtained from commercial sources, reductive amination using titanium (IV) isopropoxide and sodium borohydride yielded the target compound **2.23** in an 82% yield (Scheme 2.13).

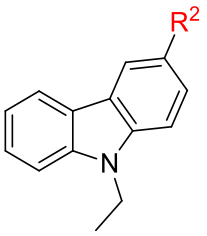


2.23, 82%

Scheme 2.13: Synthesis of compound **2.23** via reductive amination. (a) $HNMe_2$, $Ti(O^iPr)_4$, Et_3N, EtOH/DCM, rt 18 h then $NaBH_4$, rt, 8 h.

Compound **2.23** was assayed as a benzylic amine-HCl salt to aid aqueous solubility. Its effect on the p53-Y220C DNA-binding domain was assessed using differential scanning fluorimetry (DSF) and the dissociation constant was determined using isothermal scanning calorimetry (ITC) (Table 2.4). Correlation between the two assays was perfect

(-1). The substitution pattern of the solvent exposed amine had a minor effect on affinity, with the dimethylamine resulting in a 30% drop in affinity.



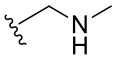
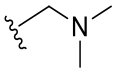
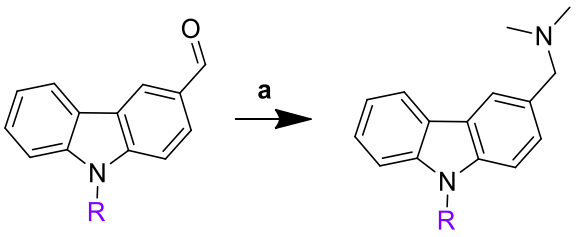
Compound	R^2	DSF ΔT_m at 125 µM	ITC K_D (µM)
PK083		0.8	124
PK8017 (2.23)		0.6	169

Table 2.4: DSF and ITC data for compounds PK083 and **2.23**. DSF and ITC assay conditions are described in Chapter 6. Correlation between DSF stabilization and ITC dissociation constant = -1 (perfect negative).

Despite the slight drop in affinity, subsequent compounds were synthesised with dimethylated benzylic amines owing to the improved yield over reductive amination with methylamine, which is much less nucleophilic. Crystallographic rationalization for this reduction in affinity is provided in the co-crystal structure of MB065, which also possess a dimethylated benzylic amine.

2.7.3 Synthesis and SAR of Compounds **2.24-2.27**

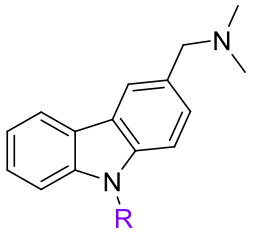
Analogous to Section 2.7.2, compounds **2.24-2.27** were synthesised in parallel from commercially available N-functionalized 3-aldehyde-substituted carbazoles via reductive amination with titanium (IV) isopropoxide and sodium borohydride (Scheme 2.14) in moderate yields (20-50%, Table 2.5).[367]



Scheme 2.14: Synthesis of compounds **2.24-2.27**. (a) $HNMe_2$, $Ti(O^iPr)_4$, Et_3N, EtOH/DCM, rt 18 h then $NaBH_4$, rt, 8 h.

Compound	R	Yield (%)
2.24	Me	20
2.25	Pr	29
2.26	iPr	27
2.27	Bn	54

Table 2.5: Yields of reductive amination for compounds **2.24-2.27**.



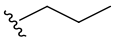
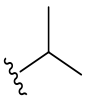
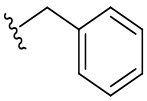
Compound	R	DSF ΔT_m [250 µM]	HSQC K_d (µM)
PK8017 (2.23)		0.96	197
PK8018 (2.24)		0.29	1600
PK8021 (2.25)		0.87	250
PK8020 (2.26)		0.25	2500
PK8016 (2.27)		0.49	/

Table 2.6: DSF and HSQC data for compounds **2.24-2.27** compared with **2.23**. DSF and HSQC assay conditions are described in Chapter 6. Correlation between DSF stabilization and HSQC dissociation constant = -0.95.

Compounds were tested for their effect on p53-Y220C DBD using DSF and their dissociation constants were determined by $^{15}N/^{1}H$ HSQC NMR (Table 2.6), an example of which is shown in Figure 2.21. Correlation between the two assays was excellent, showing a negative correlation of -0.95.

The SAR reveals that substitution of R has a dramatic effect on affinity and hence stabilization. When R < Et, there is a dramatic reduction in affinity which is due to a much worse fit of the hydrophobic anchor in the mutant induced pocket. When R > Et, a reduction in affinity is observed of the order *n*Pr < *i*Pr. From the x-ray structure of PK083 bound to p53-Y220C, N-Et does not fully occupy the Cys220 sub-pocket and hence substitution for N-Pr was expected to lead to an increase in affinity arising from a better fit of the hydrophobic sub-pocket. However, the introduction of an extra rotatable bond leads to a greater degree of entropy, hence any modest increase in affinity due to a better fit of the hydrophobic pocket may be outweighed by the entropic penalty of an extra rotatable bond, leading to a decrease in affinity.

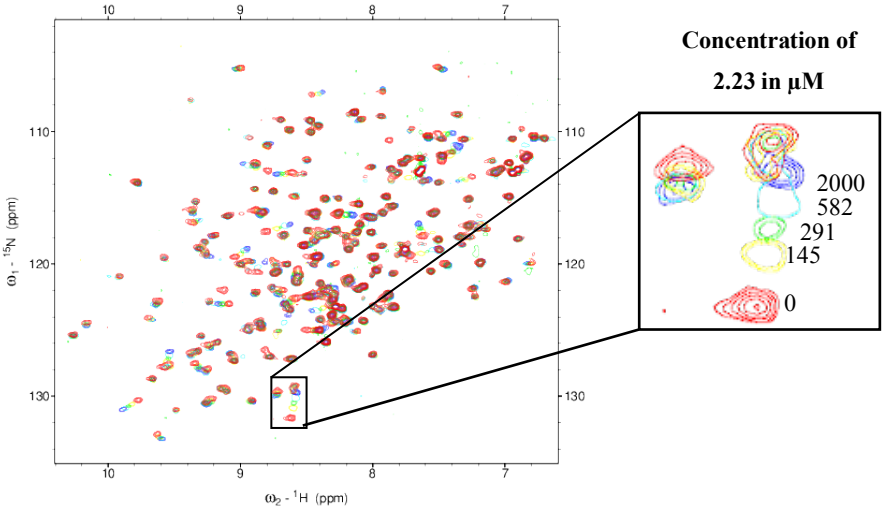
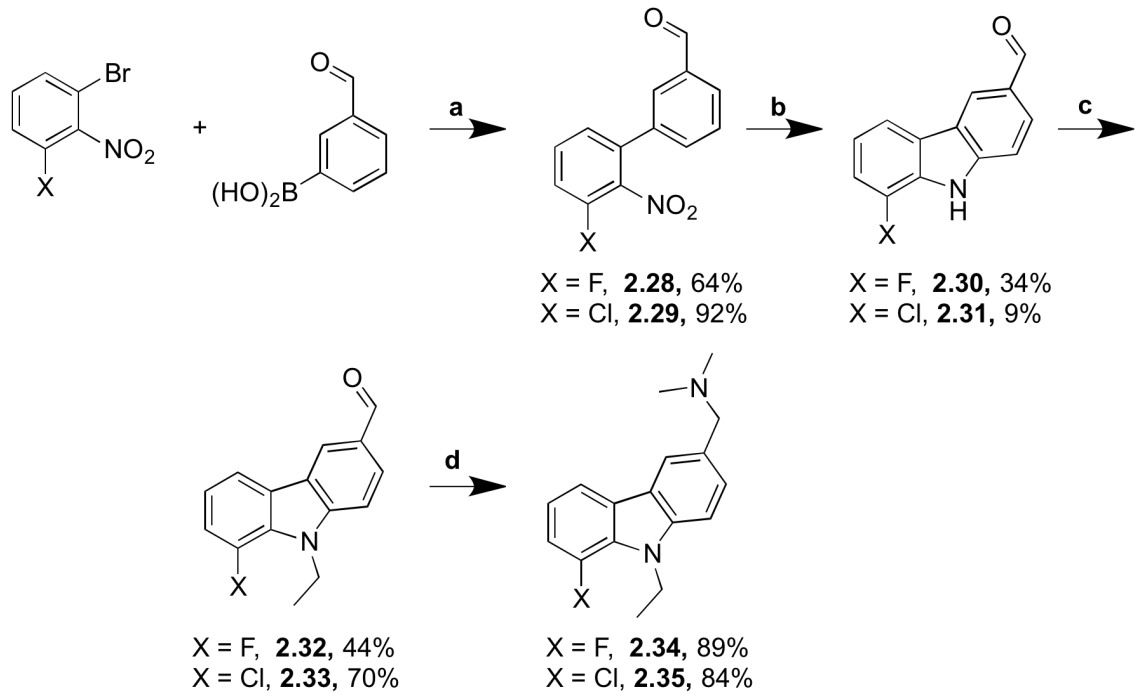


Figure 2.21: Overlay of $^{15}N/^{1}H$ HSQC NMR spectra of p53-Y220C without ligand (red) and increasing concentrations of **2.23** (145 µM, yellow; 291 µM, green; 582 µM, cyan; 2000 µM, blue).

The SAR around the size of substituent R revealed that if R > Et, a reduction in affinity occurs, hence this limitation was considered when designing all subsequent compounds.

2.7.4 Cys220 Targeting Compounds

Synthesis of **2.34** and **2.35** according the protocol described in Scheme 2.15 afforded the desired compounds. Suzuki-Miyaura cross-coupling of the appropriate aryl bromide with 3-formylphenyl boronic acid afforded intermediates **2.28** and **2.29**. Initial attempts at the Cadogan cyclization using conditions described by Freeman *et al.* failed to afford intermediates **2.30** and **2.31**, however, by the use of microwave-mediated synthesis these intermediates were accessed.[368] A nucleophilic substitution reaction with ethyl iodide using sodium hydride as a base afforded compounds **2.32** and **2.33** which, upon reductive amination yielded final products **2.34** and **2.35**.



Scheme 2.15: Synthesis of **2.34** and **2.35**. (a) $Pd(dppf)Cl_2.CH_2Cl_2$ (5 mol%), K_2CO_3, toluene, H_2O, 90 °C; (b) PPh_3, DMA, µwave, 200 °C, 2h; (c) NaH, THF, rt, 10 mins then CH_3CH_2I, 18 h; (d) $HNMe_2$, $NaBH(OAc)_3$, THF, rt, 18 h.

The failure of the Cadogan cyclization under conventional heating is owing to the de-activating effect of a powerful electron-withdrawing group *ortho* to the nitrogen. By accessing much higher reaction temperatures, this effect is overcome yielding the desired intermediate in modest yields.

The effect of compounds **2.34** and **2.35** on p53-Y220C DBD was tested using DSF. The results are summarized in Table 2.7.

Compound	R^1	DSF ΔT_m[250 µM]
PK8017 (2.23)	H	1.1
PK9335 (2.35)	Cl	0.3
PK9336 (2.34)	F	0.4

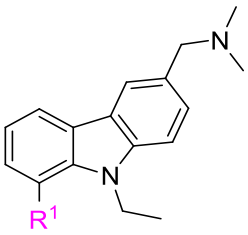
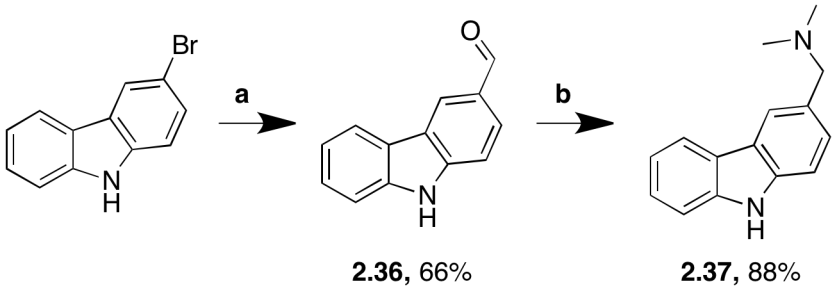


Table 2.7: DSF data for **2.23** compared with **2.35** and **2.34**. DSF assay conditions are described in Chapter 6.

Compounds **2.34** and **2.35** both display a reduction in stabilization on the p53-Y220C DBD over the original hit PK083 indicating that this position is not optimal for targeting Cys220. Owing to the very small volume of available space between Cys220 and the 1-position of the aromatic ring, attempts to target cysteine through this position were abandoned.

2.7.5 Synthesis and SAR of N-alkylated Cys220 Targeting Compounds

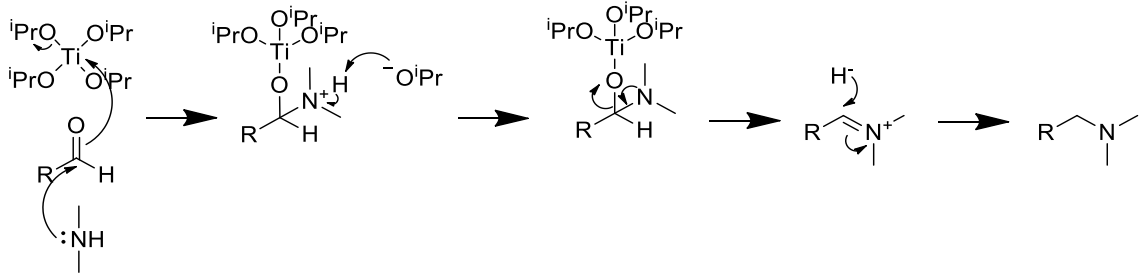
Synthesis of intermediate **2.37**, using the procedure outlined in Scheme 2.16, yielded **2.37** in an overall yield of 58%.



2.36, 66% **2.37,** 88%

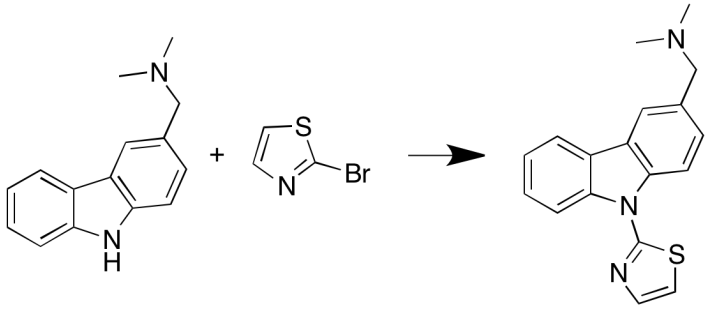
Scheme 2.16: Synthesis of privileged intermediate **2.37**. (a) NaH, THF, rt then -78 °C, *n*BuLi, -78 - -35 °C then -78 °C, DMF, THF, rt; (b) $HNMe_2$, $NaBH(OAc)_3$, THF, rt.

Reductive amination using titanium (IV) isopropoxide and sodium borohydride was unsuccessful when the carbazole contained a free N-H. By the use of sodium triacetoxyborohydride as a reducing agent this problem was circumvented.[369] The failure of the titanium method to successfully carry out the reaction is suspected to be due to the fact that carbazole can form complexes with titanium and hence is sequestered from activating the carbonyl species (Scheme 2.17).[370]



Scheme 2.17: Mechanism for reductive amination with titanium (IV) isopropoxide.

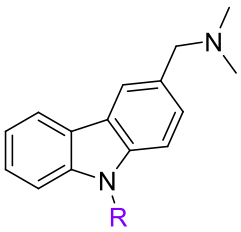
Reaction of the intermediate **2.37** using microwave-mediated Ullmann-coupling successfully yielded compound **2.38** (Scheme 2.18).[365]



2.37 **2.38,** 53%

Scheme 2.18: Synthesis of compound **2.38** by a microwave-assisted Ullmann-coupling reaction. (a) CuI (10 mol%), Cs_2CO_3, DMF, μwave, 220 °C, 40 min.

The stabilizing effect of compound **2.38** as well as PK9277, (B.S) and MB016, (M.B) on p53-Y220C was tested by differential scanning fluorimetry and are summarized in Table 2.8.



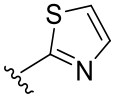
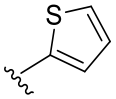
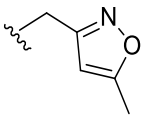
Compound	R	DSF ΔT_m [250 µM]
PK8017 (**2.23**)		0.6
PK9334 (**2.38**)		0.7
MB016		1.0
PK9227		0.5

Table 2.8: DSF data for **2.23** and **2.38** as well as MB016 and PK9227. DSF assay conditions are described in Chapter 6.

Concentration[µM]	DSF ΔT_m
1000	0.7
500	1.1
250	1.0
125	0.7
62.5	0.3

Table 2.9: DSF stabilization of MB016 on p53-Y220C DBD and different concentrations.

Substitution for N-Et by small rigid heterocyclic groups in PK9334 and MB016 lead to a minor increase in stabilization as identified by DSF, however, introduction of a benzylic heterocyclic group lead to a reduction in affinity, presumably due to the increased steric bulk and introduction of an additional rotatable bond. The more lipophilic heterocycle (N-thiophene) in MB016 outperformed PK9334 (N-thiazole) presumably due to a difference in desolation penalty. However, MB016 did not appear to show concentration dependence which could be due to poor solubility or to interference with the assay (Table 2.9). The subsequent validation of binding as well as design and synthesis of compounds where R = heterocycle was carried out by our collaborators at the LMB.

2.8 Conclusions

Attempts to target Cys220 and the mutational sub-pocket through non-covalent interactions did not lead to any confirmed affinity increases over the original hit PK083.

Substitution of the solvent exposed subsite I mono-methylated at R^2 with a dimethylated amine resulted in a 30% reduction in affinity. The structural reasoning for this is provided in Chapter 3.0.

Attempts to target Cys220 through halogenation of position R^1 resulted in a reduction in stabilization indicating that S···Cl or fluorine multipolar/hydrogen bonding did not occur. The space between position 1 on the carbazole ring at the bridge between the mutational sub-pocket and subsite II is small and hence the compounds **2.34** and **2.35** may cause a shift in the binding pose leading to the weakening or loss of the crucial hydrogen bond formed between the benzylic amine and the backbone carbonyl.

The assay results of compounds **2.23-2.27**, which were synthesized to establish the steric limit in the Cys220 mutational sub-pocket, indicate that the optimum group at R is ethyl. Substitution for *n*Pr and *i*Pr were tolerated but lead to a decrease in stabilization shown by DSF and $^{15}N/^{1}H$ HSQC. Furthermore, attempts to target the transient subsite III through arylation of position 9 showed an SAR that indicated a positive effect on stabilization. However, a lack of concentration dependence suggests promiscuous binding and/or solubility problems.

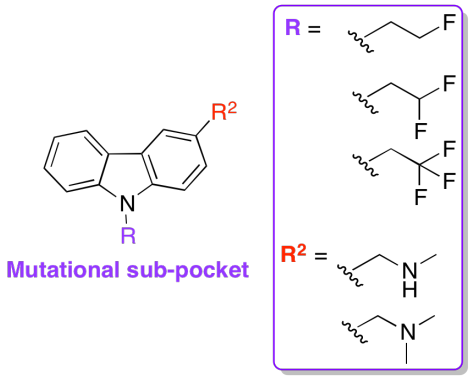


Figure 2.22: Structure of compounds designed and synthesised in Chapter 3.

Using the steric limitations derived from the SAR around the functionalization of R, Chapter 3 describes the design and synthesis of N-fluoroethylated PK083 analogues

which fall within these limitations and show increased potency towards p53-Y220C (Figure 2.22).

Chapter 3 : Harnessing Direct Fluorine-Protein Interactions in the Design of more Potent PK083 Analogues

Fluorination of compounds has been exploited extensively in the design and development of pharmaceuticals, with 20-25% of drugs on the market containing at least one fluorine atom.[371] The popularity of incorporation of fluorine into compounds is due to the predictable productive influence it has on: conformational restraint of ligands, pK_a, membrane permeability, metabolism and pharmacokinetic (PK) properties (Figure 3.3.1).[372–376] In classical medicinal chemistry, fluorine substitution is typically used at the lead optimization of the drug discovery pipeline seeking to use its ability to modulate the aforementioned properties of lead compounds towards a drug candidate with a pharmacodynamic (PD) and PK profile within acceptable safety margins.[376]

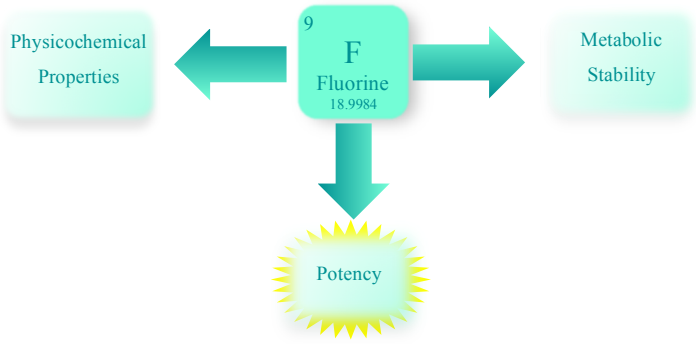


Figure 3.3.1: Key drug properties fluorine can influence.

The demonstrable versatility of fluorine within medicinal chemistry has led to numerous examples of fluorination of molecules, in which, besides the influence on the properties aforementioned, serendipitous discovery of direct fluorine-protein interactions has had a positive influence on potency.[377,378] Direct fluorine-protein interactions can be sub-divided into: polar interactions with hydrogen bond donors (e.g. amide backbone NH, polarized C_α-H from polar side chains, protein bound H_2O), hydrophobic interactions with lipophilic side chains and orthogonal multipolar interactions with a dipolar group (e.g. amide backbone C=O, amide containing side chains (Asn, Gln) and guanidinium groups (Arg) (reviewed in **2.2**). As the nature of these interactions is becoming increasingly well understood, the use of fluorination as a tool to increase a compound's potency towards its target in rational drug discovery is becoming widespread. Below, a review of the use of fluorine in medicinal chemistry is presented, concentrating on direct fluorine-protein interactions and their relative efficiency. This is followed by the

description of the design and synthesis of a more potent fluorinated PK083 analogue in which affinity gains are attributed to direct fluorine-protein interactions.

3.1 The Use of Fluorine in Medicinal Chemistry

Modern medicinal chemistry is based on a multiparameter optimization (MPO) approach, where physicochemical properties, compound safety and pharmacokinetic parameters such as: adsorption, distribution, metabolism, and excretion, (ADME) are considered in equal measures alongside potency and selectivity. Consequently, methods of modulating compound properties are crucial to medicinal chemistry. Commonly, this is accomplished by isosteric substitution at strategic positions in lead compounds, seeking to affect the physicochemical properties that are linked with structure-property relationships related to poor compound PK. For example, a large number of drugs on the market or in development contain at least one or more basic nitrogen atoms.[379] The ionization state of the nitrogen atom is not only important for potency, but affects lipophilicity, membrane permeation, and amphiphilicity.[380] It also carries a potential liability for drug-induced phospholipidosis (intracellular accumulation of phospholipids with lamellar bodies, most likely from an impaired phospholipid metabolism of the lysosome)[381] and interference with human ether-a-go-go-related gene (hERG), potassium ion channels and cardiovascular toxicity.[382–384] Modulation of the basicity of the amine by proximal fluorination, using well-developed predictive rules and structure-property relationships, provides a rational method to achieve the right balance between potency, selectivity and toxicity and is a strategy commonly utilized.[376]

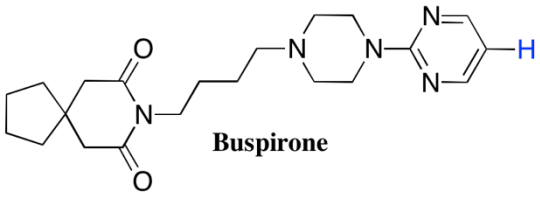
Numerous examples exist where fluorine is employed as a hydrogen isostere to modulate compound properties that show comparable potency even though its Van der Waals radii is more similar to oxygen (Table 3.1)[385] Spatially, the trifluoromethyl substitution is likened to anything from an *iso*-propyl group to a phenyl or *tert*-butyl group depending on the method used to estimate the volume.[386,387] For example, rotational energy studies on *o,o'*-disubstituted biphenyl derivatives suggest that the CF_3 group occupies a volume slightly larger than *i*Pr.[388,389] Conversely, x-ray data estimates the volume to be closer to an *iso*-butyl group.[390,391] Regardless of the ambiguity in the isosteric mimicry that fluorine substitution has on steric bulk, it has a profound effect on the electronic properties of a compound owing to its high electronegativity.[392]

Atom	Van der Waals radii (Å)	Pauling Electronegativity (χ_P)
C	1.70	2.55
H	1.20	2.20
F	1.47	3.98
O	1.52	3.44
N	1.55	3.04
Cl	1.75	3.16

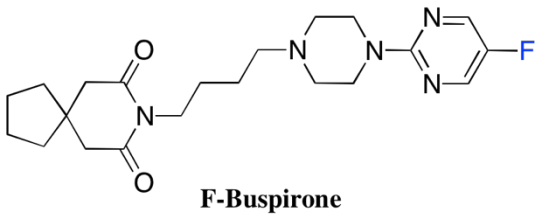
Table 3.1: Van der Waals radii and Pauling electronegativities.

3.1.1 Metabolic Stability

One of the major incentives for fluorine substitution in medicinal chemistry is the ability to enhance metabolic stability. Phase I drug metabolism by hepatic cytochrome P450 monooxygenases, a group of heme-thiolate proteins, constitutes a common mechanism of drug clearance by decreasing lipophilicity through oxidation. Metabolism by cytochrome P450's is one of the most common mechanisms of limiting compound efficacy. Substitution of a hydrogen atom in a C-H bond for a fluorine enhances metabolic stability by increasing the dissociation constant of the bond from 98.8 kcal/mol (C-H) to 105.4 kcal/mol (C-F), blocking metabolically labile moieties.[393] For example, fluorination of buspirone at the 5' position of the pyrimidine, which is commonly metabolized by CYP3A4 by hydroxylation, increases the *in vitro* $t_{1/2}$ from 4.6 min to 52 min (Figure 3.2). [394]



Buspirone



F-Buspirone

Compound	5-HT1A IC_{50} (nM)	CYP3A4 $t_{1/2}$ (min)
Buspirone	25	4.6
F-Buspirone	63	52.3

Figure 3.2: Fluorination of buspirone at the 5'- pyrimidine position blocks metabolism and increases in *vivo* $t_{1/2}$.

3.1.2 Physicochemical Properties

The physicochemical properties of drugs, such as partition coefficient (clogP and logD) and ionization constant (pK_a), affect drug PK (ADME), toxicity, and efficacy. Predictive rules, seeking to generalize the physicochemical and structural properties of orally bioavailable drugs, such as Lipinski's 'rule of 5,' provide medicinal chemists with helpful guidelines to develop drugs that have adequate human PK and safety after oral dosing and intestinal absorption (Figure 3.3).[395] Consequently, there has been much interest in methods with which to modulate these properties to allow rational design of drugs that falls within these guidelines.

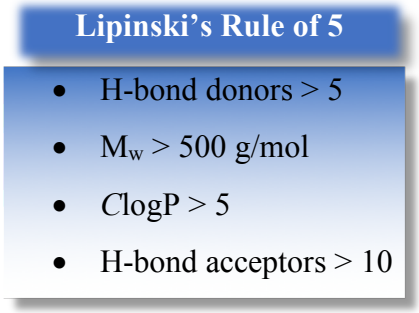


Figure 3.3: Lipinski's Rule of 5.

3.1.2.1 Modulation of pKa

The majority of drugs contain an ionizable group (Figure 3.4).[396] The ionization state of a drug compound has a profound effect on efficacy, PK and toxicity. Changes in the ionization state of a compound can lead to positive (e.g. enhance the strength of interactions) or negative (e.g. increase the desolvation penalty) effects on potency. Furthermore, the pK_a can affect solubility and permeability. Solubility is determined by both the intrinsic solubility of the neutral molecule and of the ionized species, of which the latter is much greater. Conversely, ionized molecules are less permeable than neutral molecules, hence, there is a need to strike the right balance between solubility and permeation to achieve drugs with good oral bioavailability. Broadly, the modulation of the pK_a of ionizable groups is accomplished through the introduction of electron-withdrawing groups or electron-donating groups in close proximity to the acid or base.

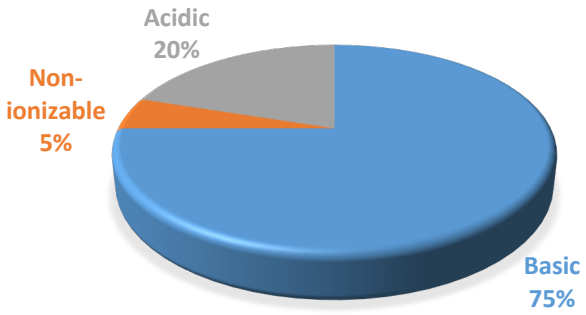
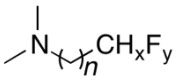


Figure 3.4: Proportion of drugs with ionizable groups.[396]

The powerful electron-withdrawing nature of fluorine affects the basicity or acidity of neighboring functional groups inductively, providing a method of pK_a attenuation. Fluorine substitution usually increases the acidity of acids and decreases the basicity of bases (Table 3.2).[393,397] The effect of sequential fluorination on linear aliphatic amines at the same carbon is approximately additive, where each fluorine substitution results in a reduction of pK_a by 1.6 to 1.7 units. The effect is very powerful to the extent that 2,2,2-trifluoromethylethylamine has a pK_a of 5.7, which, is so weakly basic that the moiety has been used as an amide isostere (Table 3.2).[398] The effect decreases exponentially as carbon chain homologation increases the distance between a fluorine atom and the amine (Table 3.3).[376]

Acid	pK_a	ΔpK_a		Amine	pK_a	ΔpK_a	
CH_3CO_2H	4.8			$CH_3CH_2NH_2$	10.7		
CH_2FCO_2H	2.6	-2.2	Acidity	$CH_2FCH_2NH_2$	9.0	-1.7	Basicity
CHF_2CO_2H	1.3	-1.3		$CHF_2CH_2NH_2$	7.3	-1.7	
CF_3CO_2H	0.5	-0.8		$CF_3CH_2NH_2$	5.7	-1.6	

Table 3.2: Effect of fluorine substitution on the pK_a of acids and bases.



n	Position	ΔpK_a
1	β - F	-1.7
2	γ - F	-0.7
3	δ - F	-0.3
4	ε - F	-0.1

Table 3.3: Effect of carbon chain homologation on modulation of pK_a by fluorine.

Fluorination of aromatic and heteroaromatic rings is a strategy primarily adopted to block metabolism (Figure 3.2), but it may also modulate pK_a. The modulating power of fluorination of aromatic rings depends on an inductive effect (σ_F), and a resonance effect (σ_R), that is represented quantitatively by a net para-effect (σ_P). Isosteric substitution of (ar)C-H by (ar)C-F has a high electron-withdrawing inductive effect (σ_F = 0.45). However, this is substantially attenuated by a large resonance component (σ_R = -0.39) leading to a weak net para-effect (σ_P = 0.06). Substitution for CF_3 leads to a much more powerful modulating effect on substituents as π-delocalization is not possible (Table 3.4).[375]

Substituent	Para-effect (σ_P)	Inductive effect (σ_F)	Resonance Effect (σ_R)
F	0.06	0.45	-0.39
CF_3	0.54	0.38	0.16

Table 3.4: Substituent Hammet constants.

3.1.2.2 Modulation of Lipophilicity

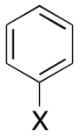
Analogously to pK_a, the lipophilicity of a compound has a considerable effect on compound PK, toxicity and potency dominantly through affecting solubility and permeability. Permeability is typically measured as the rate at which a molecule is passively diffused or actively transported (uptake/efflux) across membranes (x 10^{-6} cm/s). Passive diffusion is more common than active transport. Orally administered drugs adsorbed through passive diffusion must be small enough and lipophilic enough to pass through a cell's apical lipid bilayer membrane but not so lipophilic that they get trapped. The lipophilicity of a drug can be expressed through its partition coefficient (log*P*) between octanol and water. Strategic fluorination of lead compounds is a common method used to influence permeation and solubility by modulating lipophilicity, association with pendent H-bond donors, or reduction of amine basicity.

Most orally administered drugs have a log*P* of between 1-5. A common misconception is that fluorination of compounds always increases lipophilicity. Fluorination of aliphatic hydrocarbons decreases the lipophilicity owing to the highly dipolar nature of the C-F bond (Table 3.5).[399] Conversely, introduction of aromatic or vinyl fluorine typically

increases lipophilicity where differences arise from resonance electron donation for unsaturated fluorination (Table 3.6).[399]

Compound	Log*P*
CH_3CH_3	1.81
CH_3CHF_2	0.75

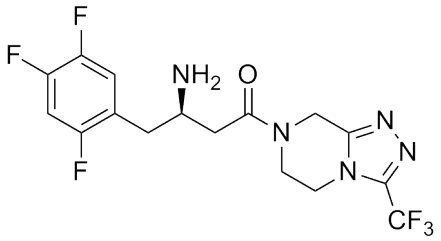
Table 3.5: Effect of fluorination on log*P* on aliphatic hydrocarbon groups.



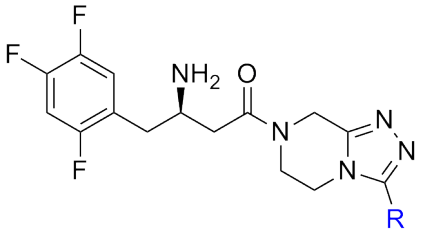
X	Π_x
CH_3	0.56
CF_3	0.88

Table 3.6: Effect of fluorination on Hansch-Leo hydrophobic parameter of aromatic groups. $\Pi_X = \log (P_x/P_H)$ (octanol/H_2O).

There are numerous examples of successful fluorination introduction strategies to improve permeation through modulating lipophilicity, associating with H-bond donors or reducing basicity of amines. For example, during the development of the oral antihyperglycemic dipeptidyl peptidase-IV (DPP) inhibitor, Sitagliptin, trifluoromethylation of the triazole moiety resulted in a marked improvement in oral bioavailability (F) (Table 3.7).[400]



Sitagliptin

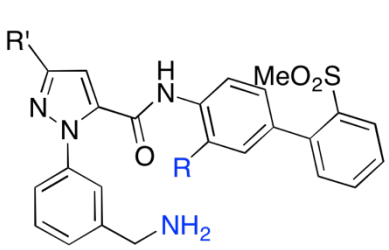


R	Clp (mL/min/kg)	$t_{1/2}$ (h)	F (%)
CH_3CH_2	70	1.7	2
CF_3 (Sitagliptin)	60	1.7	76

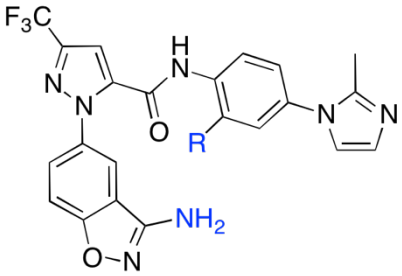
Table 3.7: Fluorination of DPP-IV inhibitors increases oral bioavailability (F).

Furthermore, increases in permeability were observed in the closely related factor Xa inhbitors **3.1** and **3.2** through isosteric substitution of hydrogen to fluorine ortho- to the anilide NH (Table 3.8).[401] The increase in permeability observed for fluoro-susbstituted compounds **3.1b, d** and **3.2b** compared to their matched pairs **3.1a,** c and **3.2a** may be due to an electrostatic interaction between the fluorine atom and the pendant N-H that effectively masks the hydrogen bond donor capability, enhancing permeation. This is

further supported by the observation that the ortho-nitrile of compound **3.2c** did not show enhanced permeability.



3.1



3.2

Compound	R	R'	Caco-2 permeability (x 10^{-6} cm/s)
3.1a	H	CH_3	1.20 ± 0.09
3.1b	F	CH_3	3.14 ± 0.10
3.1c	H	CF_3	3.38 ± 0.08
3.1d	F	CF_3	4.86 ± 0.33
3.2a	H	n/a	0.8
3.2b	F	n/a	7.4
3.2c	CN	n/a	<0.1

Table 3.8: Caco-2 permeability of factor Xa inhibitors **3.1a-d** and **3.2a-c**. A potential electrostatic interaction is formed between the pendant amine (blue) and R (blue) when R=F leading to an increase in potency.

3.1.3 Conformational Control

Beyond fluorine's electron-withdrawing influence on compound properties, the introduction of aliphatic fluorine into compounds has the ability to increase potency and selectivity through the stabilization of preferential ligand-binding conformations.[321] The high electronegativity of fluorine results in a highly-polarized C-F bond that presents both a strong dipole moment and a low-lying σ^*_{CF} orbital available for hyper-conjugative donation. A combination of these effects results in a strong preference for vicinal functionality to align *gauche* to fluorine, where the dominant underlying contributions to the stabilizing interactions that constitute the *gauche* anti-periplanar conformation depends on the nature of the vicinal functionality X (Figure 3.5, Table 3.9).[402] Didactically, this may be considered a donor-acceptor interaction closely related to the anomeric effect in carbohydrates (Figure 3.5). [403–405]

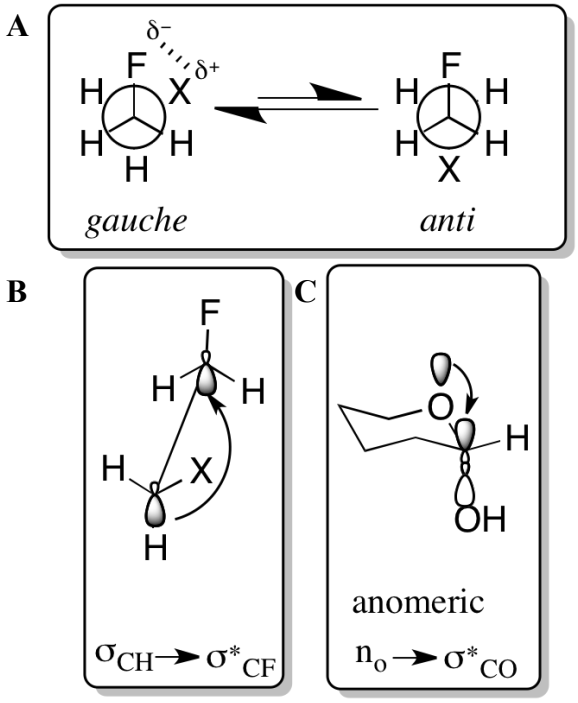


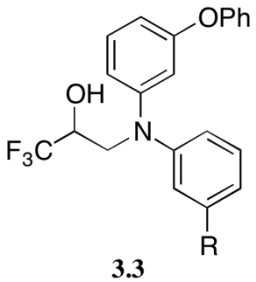
Figure 3.5: The *gauche* effect. (A) Newman projections of *gauche* and *anti* conformations. Additional stabilization where X is an electron deficient heteroatom (NH_3^+) shown by dashed line between partial charges. (B) Sawhorse representation showing molecular orbitals of the hyperconjugative component of the *gauche* effect showing donation from donor HOMO (σ_{CH}) to acceptor LUMO (σ_{CF}). (C) Molecular orbital representation of the analogous donor-acceptor anomeric effect in carbohydrates.

X	Stabilization Energy (kcal/mol)	Underlying interaction
F	0.5-1.0	$\sigma_{CH} \rightarrow \sigma^*_{CF}$ hyperconjugative donation
OH	1.0-2.0	$\sigma_{CH} \rightarrow \sigma^*_{CF}$ hyperconjugative donation and C-F···O-H electrostatic interaction
NH_2	0.9-1.0	Intramolecular C-F···H-N hydrogen bonding
NH_3^+	5.8	Electrostatic interaction between $CF\delta^{-}\cdots\delta^{+}$ NH_3^+dipoles
OAc	1.6	$\sigma_{AcO\text{-}H} \rightarrow \sigma^*_{CF}$ hyperconjugative donation
NHAc	1.8	Electrostatic interaction between $F\delta^{-}\cdots\delta^{+}$NH dipoles

Table 3.9: Calculated energy differences between *gauche* and *anti*-isomers of selected X-substituted fluoroethane derivatives shown in Figure 3.5A.[402]

Utilization of the *gauche* effect towards enhancing potency and selectivity by restricting ligand conformation has been displayed in the design of superior inhibitors of cholesteryl ester transfer protein (CETP) towards a treatment to reduce the risk of atherosclerosis by improving lipid blood levels.[406] Isosteric substitution of the 3-ethoxy moiety (**3.3a**) for a

3-tetrafluoroethyoxy substituent (**3.3b**) led to a 8-fold improvement in potency (Table 3.10).



3.3

Compound	R	IC_{50} (μM)
3.3a	OCH_2CH_3	1.6
3.3b	OCF_2CF_2H	0.2

Table 3.10: IC_{50} of compounds **3.3** against CETP.

Molecular modelling suggested that the main reason behind the improvement in potency for **3.3b** was due to the tetrafluoroethyl groups preference to adopt an out of plane orientation with respect to the phenyl ring (calculated Ar-O-CF_2 bond angle to be perpendicular ~ 90 º) in contrast to the ethoxy group, which adopts a more co-planar conformation (Figure 3.6).[406]

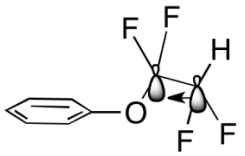


Figure 3.6: Out of plane orientation of tetrafluoroethyl group as calculated by Massa *et al*.[406]

3.1.4 Direct Fluorine-Protein Interactions

The use of fluorine to achieve desirable PK is widespread because of the well understood effect that fluorination has on ADME properties. In contrast, most examples of direct fluorine-protein interactions that improve potency are discovered serendipitously. Comparatively, the use of fluorination to improve potency through direct fluorine-protein interactions is far less well understood, resulting in a dearth of examples that use fluorination strategies in rational compound design for improving potency. Over the last decade, many systematic scans of protein-ligand structures (PDB) and small-molecule structures (CSD) that display F···R (R = protein or small molecule) close contacts have sought to statistically determine the nature and strength of these interactions, seeking to define a set of rules in order to utilize fluorination strategies towards improving ligand-protein affinities.[322,323,407,408] However, these studies have led to confusion arising from

differing conclusions made regarding fluorine's ability to act as a hydrogen bond acceptor. For example, a highly-cited review by Dunitz and Taylor concluded that organofluorine hardly ever accepts hydrogen bonds,[407] whereas, Bissantz *et al*. describes organofluorine as frequently forming weak hydrogen bonds with XH, where X = O, N.[326] Consequently, a comprehensive set of rules that define how fluorine forms intermolecular interactions remains elusive. In deference to this ambiguity, these interactions will be referred to 'polar interactions with hydrogen bond donors.' Broadly, direct fluorine-protein interactions may be classified either hydrophobic interactions with lipophilic side chains, polar interactions with hydrogen bond donors or orthogonal multipolar interactions with a dipolar group (See 2.2 for a review on multipolar interactions).

3.1.4.1 Hydrophobic Interactions of Fluorine

The IUPAC definition of hydrophobic interactions is as follows:

> 'the tendency of hydrocarbons (or of lipophilic hydrocarbon-like group in solutes) to form intermolecular aggregates in an aqueous medium, and the analogous intramolecular interaction.'[409]

It is the driving force behind a plethora of effects, such as: why an oil/water mixture spontaneously separates, why soluble proteins fold with a hydrophobic core and a hydrophilic outer surface, and importantly, why small molecules associate in protein binding pockets with mutual burial of hydrophobic surfaces.[410–412] The hydrophobic effect can be visualized by considering the entropy of ligand-protein association. A hydrophobic solute (ligand, hydrophobic pocket of protein) disrupts the structure of bulk water and decreases entropy because of stronger bonding and ordering of water molecules around the solute. These disruptions are minimized upon association of nonpolar solute molecules (ligand-protein association). Water then constitutes one larger 'cage' structure around the combined solutes, whose area is smaller than the combined surface areas of isolated solutes. This maximizes the amount of free water and thus the entropy (Figure 3.7). Hence, the hydrophobic effect is commonly attributed to the large positive entropic term associated with the displacement of water molecules arranged around the hydrophobic surface.

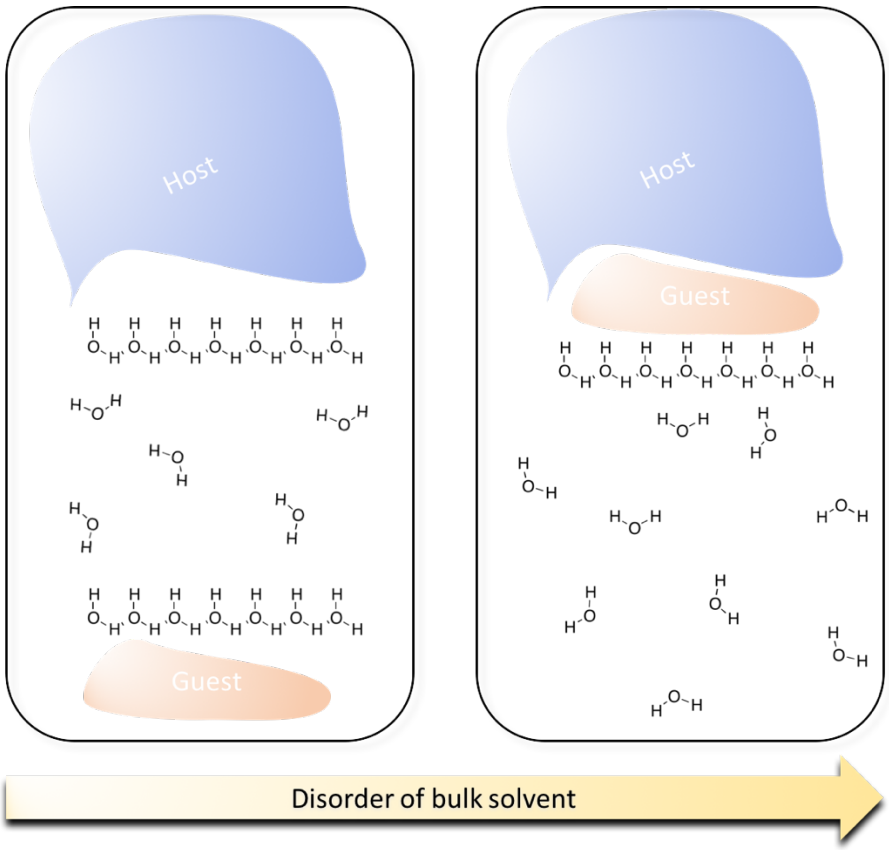
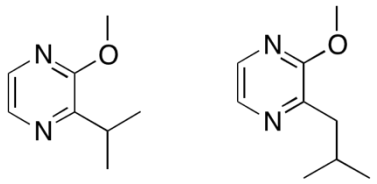


Figure 3.7: The hydrophobic effect explained by an increase in entropy. Figure shows the increased disorder of bulk solvent on the association of a guest (ligand), with a host (hydrophobic pocket of target protein). For simplicity, only one shell of solvation around the hydrophobic solutes is depicted.

However, recent studies challenge this simplistic explanation, demonstrating that the hydrophobic effect can originate from enthalpy gains as well.[413–418] Spectroscopic evidence has noted that the hydrogen bonds at hydrophobic surfaces are weaker and hence the water molecules are more flexible than presumed using the entropic model of the hydrophobic effect.[419,420] Complexation thermodynamics that are driven by enthalpy are termed the 'non-classical hydrophobic effect.'[421] Homans studied the complexation thermodynamics of inhibitors of the mouse major urinary protein (MUP) using a combination of ITC, x-ray crystallography and ^{15}N and methyl side-chain ^{2}H relaxation measurements.[415] Thermodynamics derived from ITC measurements suggested that the free energy of binding was driven by favourable enthalpic contributions (Table 3.11). $\Delta G°$ for binding is complex, involving contributions from desolvation of the ligand and binding pocket upon binding, the free energy arising from new ligand-protein interactions and, changes in the structural and dynamic properties of the protein and ligand, including loss of translational and rotational entropy ($S = k_B \ln\Omega$, natural logarithm of the number of configurations (Ω) multiplied by Boltzmann constant (k_B)).[422] Using this ITC data,

along with calculations of desolvation of the MUP ligands using vapour-water partition experiments, the contributions to the overall free binding energy were assessed according to the thermodynamic cycle described for the formation of a ligand-protein interaction in solvent water (Figure 3.8).[422–426]



IPMP **IBMP**

Compound	ΔG° (kJ/mol)	ΔH° (kJ/mol)	TΔS° (kJ/mol)	k_d (μM)
IBMP	-38.5	-47.9	-9.4	0.3
IPMP	-33.9	-44.5	-10.8	1.8

Table 3.11: ITC determination of the contributions of the overall enthalpy and entropy on the free binding energy of MUP inhibitors 2-methoxy-3-isopropylpyrazine (IPMP) and 2-methoxy-3-isobutylpyrazine (IBMP).

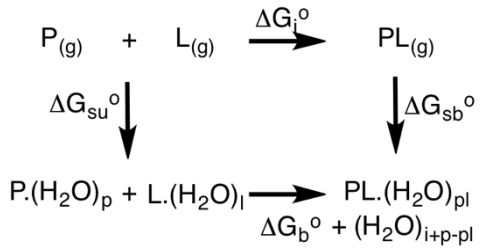


Figure 3.8: Thermodynamic cycle for a ligand-protein interaction in solvent water. $\Delta G_b°$ represents the observed free energy of binding, $\Delta G_i°$ represents the intrinsic solute-solute interaction in the absence of solvent, $\Delta Gs_b°$ and $\Delta G_{su}°$ represent the free energy of solvation of the complex and the free species respectively. Since G is a state function, the sum over the cycle is zero, so: $\Delta G_b° = \Delta G_i° + (\Delta Gs_b° - \Delta G_{su}°)$. Where $\Delta Gs_b°$ is equated with the solvation of the complex $\Delta G_{solvPL}°$ and, $\Delta G_{su}°$ comprises the sum of the solvation free energies of the free protein and ligand, hence, $\Delta G_b° = \Delta G_i° + (\Delta G_{solvPL}° - \Delta G_{solvP}° - \Delta G_{solvL}°)$. Since G is a state function, the sum over the cycle is zero so: $\Delta G_b° = \Delta G_i° + (\Delta Gs_b° - \Delta G_{su}°)$, where, $\Delta Gs_b°$ is equated with the solvation of the complex $\Delta G_{solvPL}°$ and, $\Delta G_{su}°$ comprises the sum of the solvation free energies of the free protein and ligand, hence, $\Delta G_b° = \Delta G_i° + (\Delta G_{solvPL}° - \Delta G_{solvP}° - \Delta G_{solvL}°)$.

To summarize, the major contribution to binding was attributed to the solvent driven enthalpic effect of desolvation of the protein binding pocket which arose due to the sub-optimal occupancy of water in the binding pocket (Table 3.12).[415] This was due to the occluded cleft structure of the pocket that prevents water from simultaneously occupying the pocket and forming hydrogen bonds with bulk water leading to an entropic preference to occupy the hydrophobic pocket but an enthalpic preference to form hydrogen bonds with their neighbours. Similar observations have been made for hydrophobic pockets of

other proteins which bear resemblance with the IPMP-MUP binding pocket (narrow hydrophobic cleft).[427]

Entropic contribution	Description	Value (kJ/mol)	Enthalpic contribution	Description	Value (kJ/mol)
$T\Delta S_i°$	Protein degrees of freedom	-0.8 ± 3.8	$\Delta H_i°$	New protein-ligand interactions	~ -76
	Ligand degrees of freedom	~ -37		Changes in ligand/protein structure	~ 0
$-T\Delta S_{solvL}°$	Ligand desolvation	+26.7 ± 8.4	$\Delta H_{solvL}°$	Ligand desolvation	+43.8 ± 8.2
$T\Delta S_{solvPL}°$-$T\Delta S_{solvP}°$	Desolvation of protein/complex	+0.4 ± 9.2	$\Delta H_{solvPL}°$-$\Delta H_{solvP}°$	Desolavtion of protein/complex	-12.3 ± 8.4
$T\Delta S_{obs}°$	Observed entropy	-10.7 ± 0.5	$\Delta H_{obs}°$	Observed enthalpy	-44.5 ± 0.4

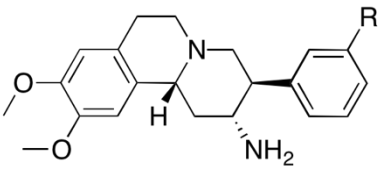
Table 3.12: Thermodynamic decomposition of binding of IPMP to MUP.

Whilst the understanding behind the classical and non-classical hydrophobic effect is far from comprehensive, it appears the key to successful productive utilization is to consider the hydration state of the unbound protein pocket to point out key hydrophobic areas that can be matched with similarly hydrophobic ligand moieties i.e.' hydrophobic hotspots'. Then, by SAR's, optimize the relevant ligand atoms/functionalities towards an optimum fit to the hydrophobic pocket. This is commonly accomplished through the introduction of hydrocarbon chains, and through fluorination. The latter owing to both the hydrophobic nature of fluorine (Section 3.1.2.2) and, the larger Van der Waals radii fluorine has compared with hydrogen.

Numerous examples of significant improvements in potency through increasing the hydrophobic surface area that have been attributed to hydrophobic interactions, many involve fluorine. For example, the fluorination of inhibitors of DPP-IV.[428] The SI specificity pocket of DPP-IV is composed of several hydrophobic side chains (Val, Trp and Tyr). Substitution of the *meta*-phenyl hydrogen for a fluoromethyl group led to a 400-fold improvement in binding affinity (Figure 3.9). An x-ray structure of **3.4** bound to DPP-IV shows the formation of five short hydrophobic contacts between the CH_2F group and the surrounding lipophilic side chains. The marked improvement in affinity of the

fluoromethylated analogue **3.4c** over the methylated analogue **3.4b** was attributed the slightly larger volume that fluorine occupies compared with hydrogen, and the commonly observed higher hydrophobic interaction strength that fluorine has respective to hydrogen.

A



3.4

Compound	R	K_d (nM)
3.4a	H	200
3.4b	CH_3	4.6
3.4c	CH_2F	0.5

B

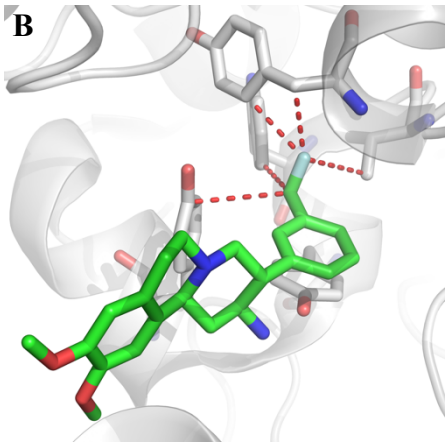


Figure 3.9: SAR of DPP-IV inhbitors **3.4a-c**. (A) Structure and K_d's of compounds **3.4ab-c.** (B) X-ray structure of **3.4c** bound to DPP-IV (PDB: 3KWJ). Short hydrophobic contacts are displayed in dashed red lines.

3.1.4.2 Polar interactions with hydrogen bond donors

Systematic scans of crystallographic databases (PDB, CSD) found polar interactions of the type C-F···H-X, where X is O, N, or C, occur frequently.[322,323,407,429–436] The nature of these interactions is still the subject of some debate, with some authors attributing it as a weak hydrogen bond and others as a multipolar interaction.[407,429,430,432,433] However, regardless of the exact nature of the interaction, they are, in general, found to contribute weakly to the free energy of binding, typically possessing interaction energies of < 3 kcal/mol.[323,437] Hence, molecular recognition and association ascribed to organofluorine forming polar interactions with hydrogen bond donors are at risk of overinterpretation.

At the core of the ambiguity seems to be whether these interactions are hydrogen bonds or not. Pauling described the hydrogen bond in his classic work of 1931, *The Nature of the Chemical Bond,* as,

> 'Only the most electronegative atoms (such as N, O, F) should form hydrogen bonds (denoted X-H···Y), and the strength of the bond should increase with increase in the electronegativity of the two bonded atoms,'

concluding that it must be electrostatic in nature and not chemical (covalent).[392] Since then, chemists have identified numerous examples of hydrogen bonding which do not fit with Pauling's definition. From Pauling's definition, fluorine, the most electronegative element, should form strong hydrogen bonds with strongly acidic protons. Indeed, anionic fluoride ions form some of the strongest hydrogen bonds $(H\text{-}F\text{-}H)^-$ and are frequently cited in textbooks as illustrative examples of hydrogen bonding.[438] However, C-F···H-X interactions are typically very weak, challenging this definition of hydrogen bonding.

Introduction of a hydrogen bond acceptor into a ligand in the rational design of more potent compounds, seeking to form a strong hydrogen bond, can lead to a smaller gain in free energy of binding than expected. Intuitively, this may be explained by several reasons, for example, the higher desolvation penalty associated with polar functional groups. Studies correlating hydrogen bond acceptor strength with proton basicity show that highly basic acceptors do not form the strongest hydrogen bonds, challenging the historic definition of increasing hydrogen bond strength with increasing electronegativity.[439] Instead, properties such as high proton affinity and high polarizability play a role in determining the strength of the acceptor. This can be used to explain the weak strength of interactions involving fluorine as an acceptor as it has a low proton affinity and low polarizability.

Recently, Desiraju *et al.* redefined the hydrogen bond as

> 'An attractive interaction between a hydrogen atom from a molecule or a molecular fragment H-X, in which X is more electronegative than H, and an atom or group of atoms in the same or different molecule, in which there is evidence for bond formation.'[440,441]

This includes the possibility that either donor or acceptor is weak or possesses moderate electronegativity. This definition amends the range of hydrogen bond stabilization energies to include the very strong hydrogen bonds (170 kJ/mol), formed between ionic acceptors and acidic donors, and the very weak hydrogen bonds(2 kJ/mol), formed by weak donors and acceptors. Desiraju differentiates the weak hydrogen bond from a Van

der Waals interaction stating that weak hydrogen bonds display directionality, whereas Van der Waals interactions do not.[442] This expands the definition to allow for the weak hydrogen bond acceptor nature of fluorine as well as providing grounds to differentiate whether an interaction can be classified as a hydrogen bond or as a Van der Waals interaction.

Despite the increasing interest in the formation of weak interactions, a predictive model for the formation of C-F···H-X remains elusive. Most studies have sought to draw statistical relevance from database searches for C-F···H-X, where the Van der Waals radius is less than the sum of the respective atoms. These searches are generally in agreement that these interactions are weak and have a subtle effect on protein-ligand recognition and association.

Vulpetti *et al*. have developed a local environment of fluorine (LEF) library which can be used to probe for fluorphillic environments. This library was constructed using a local environment to fluorine (LEF) rule of shielding.[443,444] This LEF rule was based on the observation that compounds with similar local-fluorine environments tend to have similar chemical shifts, and that these shifts are modulated by the number of oxygen, nitrogen and halogen atoms that are situated α, β, and γ with respect to fluorine. In general, the ^{19}F chemical shift correlates with the type of interaction that fluorine is prone to make with the protein, where highly shielded fluorine observed frequently in contact with hydrogen bond donors (Figure 3.10). Comparison of protein-ligand x-ray structures present in the PDB and CSD support the proposed correlation between ^{19}F chemical shift and the amphipathic character of fluorine (formation of weak hydrogen bonds, multipolar interactions or hydrophobic interactions).

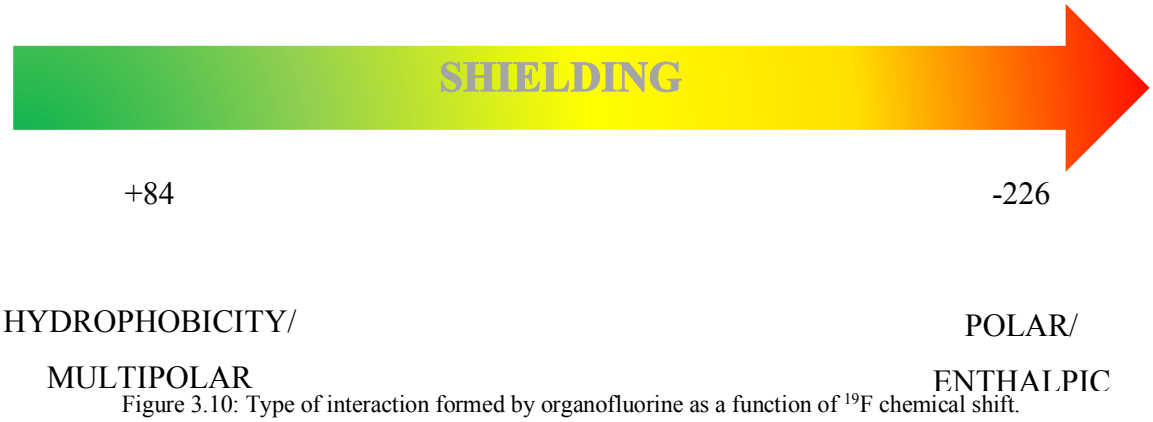


Figure 3.10: Type of interaction formed by organofluorine as a function of ^{19}F chemical shift.

Several illustrative examples of organofluorine forming polar interactions with HX are frequently cited in review articles, seeking to demonstrate how these weak interactions can be used in rational drug discovery towards modest improvements in affinity. Ortho fluorination of the aromatic ring of an sitagliptin analogue **3.5a** to form a 2,4,5-fluroinated phenyl ring (**3.5b**) results in the formation of two polar contacts with Arg125 and Asn710 NH_2 moieties (Figure 3.11).[400] Comparison with the 3,4-difluoro substituted analogue reveals that ortho-fluorination results in an improvement in potency from 455 nM (**3.5a**) to 68 nM (**3.5b**).

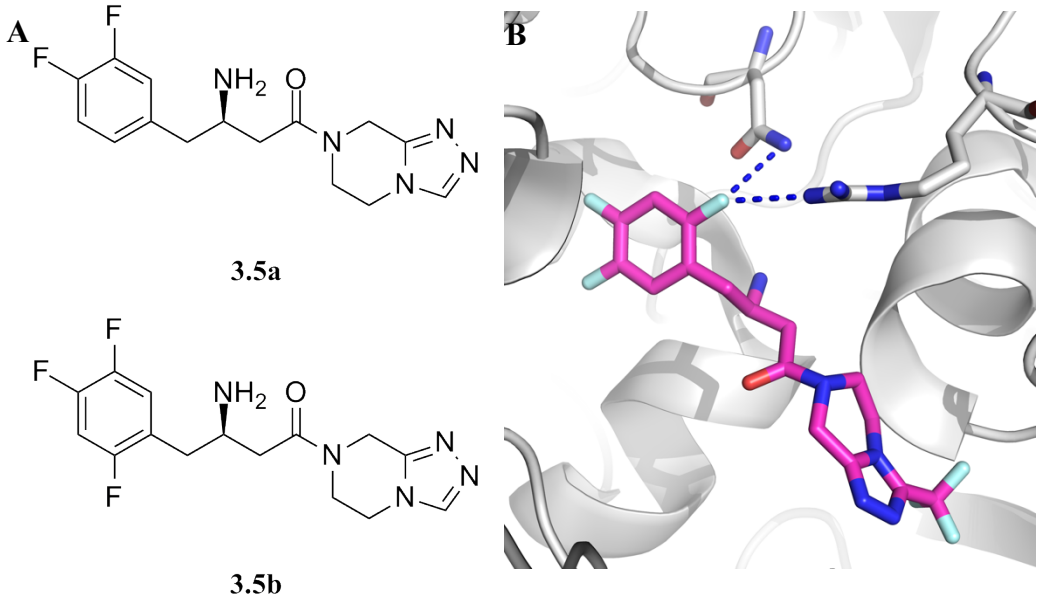
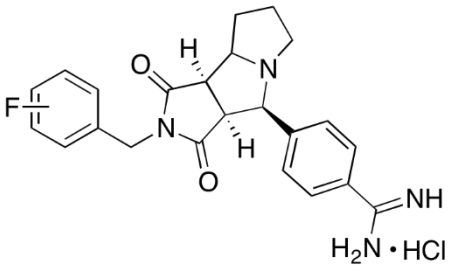


Figure 3.11: Increase in potency by fluorination of sitagliptin analogue **3.5a**. (A) Chemical structure of sitagliptin analogues **3.5a** and **3.5b**. (B) Structure of **3.5b** bound to DPP-IV (PDB: 1X70). For simplicity, just the polar interactions between the ortho-fluorine are shown (blue dashed lines).

A systematic fluorine scan of the phenyl ring of the thrombin inhibitor (Figure 3.12) revealed that the 4-substituted analogue showed increased selectivity and potency.[445] An x-ray structure of inhibitor **3.6d** bound to thrombin revealed that the fluorophenyl group forms 5 short contacts with Glu97 and Asn98 (multipolar) and Ile174 (polar interaction with hydrogen bond donor) below the Van der Waals radii. The authors reasoned that the formation of these contacts was responsible for the improved selectivity and affinity.

These illustrative examples establish two key point on these interactions. Observed polar interactions between fluorine and hydrogen bond donors are seldom isolated interactions and typically occur in conjunction with other interactions (orthogonal multipolar) in particularly fluorophilic environments. Whilst these interactions are weak in nature, modest to large improvements in affinity can be gained through targeting these fluorophillic environments through subtle fluorine-protein interactions and through indirect electronic effects.

A



3.6

Compound	F- position	K_i (μM)	Selectivity (%)
3.6a	-	0.31	15
3.6b	2	0.50	9.8
3.6c	3	0.36	26
3.6d	4	0.057	67
3.6e	2,6	0.61	9.0
3.6f	3,5	0.59	25
3.6g	1,2,3,4,5	0.27	44

B

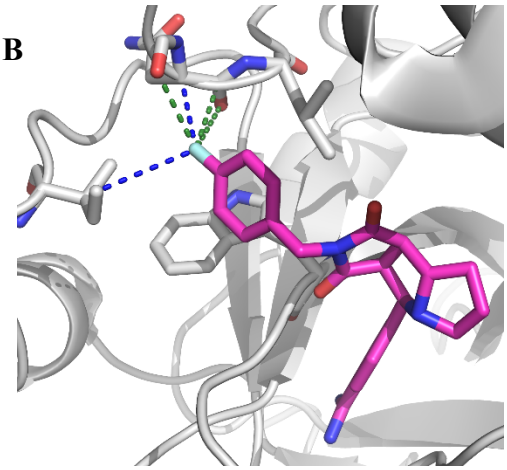


Figure 3.12: Increase in potency through a systematic fluorine scan of Thrombin inhibitor **3.6**. (A) Structure of fluorinated thrombin inhibitor and table showing activity. Uncertainty in K_i values is ± 20%. Selectivity is relative to trypsin. (B) Structure of **3.6d** bound to thrombin (PDB: 1OYT). Polar interactions with hydrogen bond donors shown in blue dashed lines. Multipolar interactions shown in green dashed lines.

3.2 Design of fluorinated PK083 analogues to target p53-Y220C

Undoubtedly, the formation of direct fluorine-protein interactions can improve potency and selectivity. However, as these interactions are weak in nature and not fully understood, targeting 'fluorophilic' pockets or sub-pockets, rather than seeking to form specific polar interactions with hydrogen bond donors, hydrophobic interactions or multipolar interactions, constitutes a more general strategy towards seeking to use direct fluorine-protein interactions towards affinity improvements.

Section 2.3 described the discovery of fragment 2.9 in which two fragments occupied the Y220C mutation site with the trifluoromethyl groups oriented toward Cys220 (Figure 2.14C).[291] This is indicative of the Cys220 mutation site being potentiallly fluorophilic. Figure 3.13 highlights this hotspot and identifies the best position of PK083 that can be fluorinated to target this potentially fluorophilic hotspot. To assess this hypothesis, DFT-D calculations were conducted‡ to computationally assess the fluorophilicity of the central mutation cavity.[446]

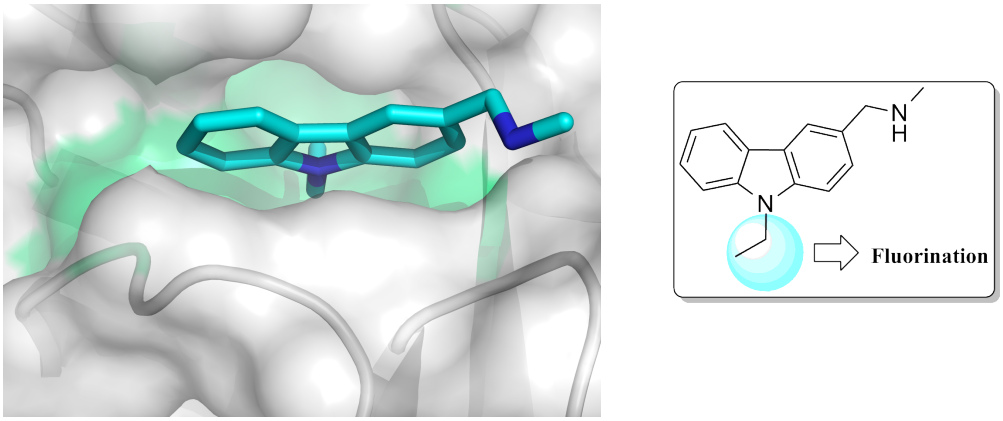
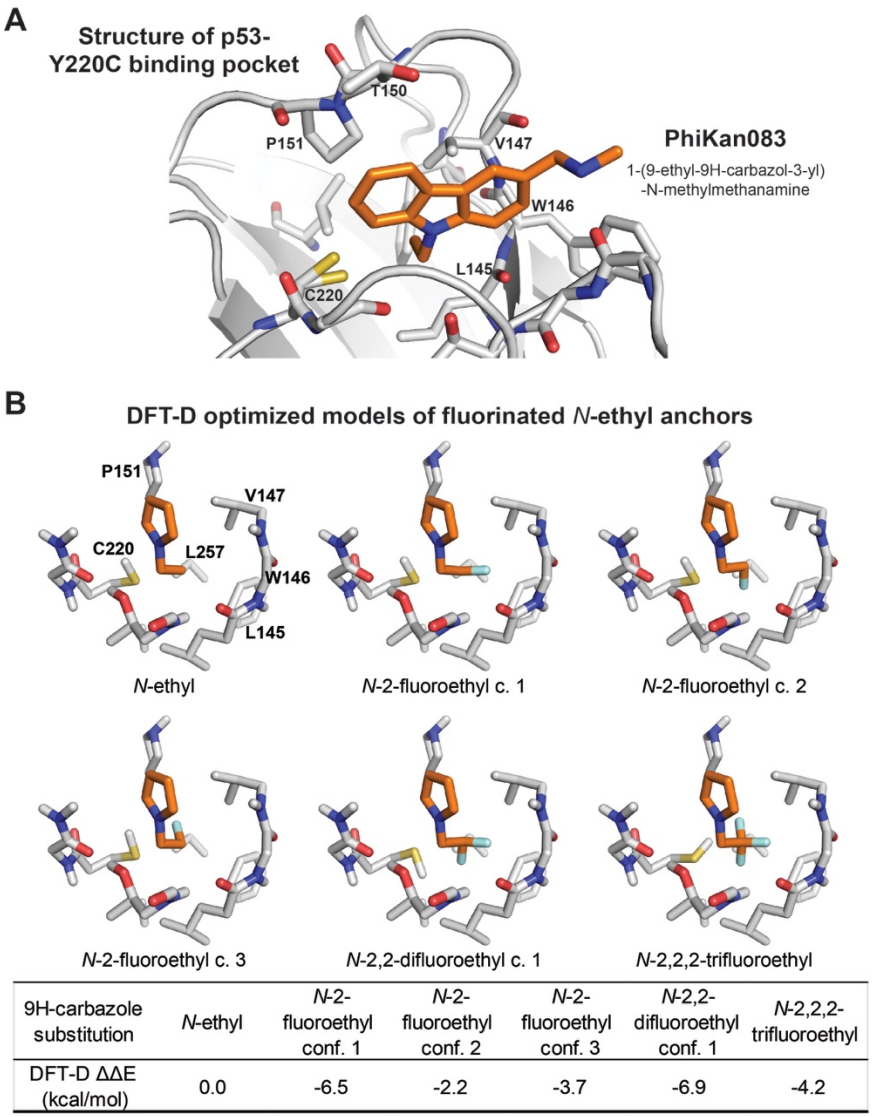


Figure 3.13: Targeting the fluorophilic hotspot of p53-Y220C. (A) Structure of PK083 bound to p53-Y220C (PDB: 2VUK), potential fluorophilic hotspot highlighted in green. (B) Chemical structure of PK083 showing the area that can be fluorinatied to target this hotspot.

3.2.1 DFT-D Calculations

In the crystal structure of p53-Y220C in complex with PK083 (Figure 3.14A), the ethyl moiety of PK083 is near the carbonyl groups of Leu145 and Trp146, and the thiol group of Cys220.[261] Given the frequent and well-characterized interactions between organofluorine groups and protein backbone amides, as well as the less frequently observed interactions between fluorine and sulfur atoms,[323] investigation as to whether gains in binding affinity could be achieved via fluorinated ethyl substituents using DFT-D calculations at BLYP-D3/def2-SVP level with a truncated model of PK083 bound to the p53-Y220C binding pocket (Figure 3.14B). Except for the sulfur atom of Cys220, all heavy atoms of the Y220C binding pocket as well as the nitrogen and C-3 atom of the pyrrole ligand model were kept frozen during the calculations.

‡ DFT-D studies were conducted by Dr. Matthias Bauer



9H-carbazole substitution	*N*-ethyl	*N*-2-fluoroethyl conf. 1	*N*-2-fluoroethyl conf. 2	*N*-2-fluoroethyl conf. 3	*N*-2,2-difluoroethyl conf. 1	*N*-2,2,2-trifluoroethyl
DFT-D ΔΔE (kcal/mol)	0.0	-6.5	-2.2	-3.7	-6.9	-4.2

Figure 3.14: Binding mode of the p53-Y220C stabilizer PK083 and fluorinated model systems (A) Experimentally determined binding mode of PK083 (orange sticks) to the mutation-induced surface crevice of the p53 mutant Y220C (PDB: 2VUK). (B) Snapshots of DFT-D optimized models of the PK083 *N*-ethyl group and its fluorinated derivatives (orange sticks) bound to the Y220C surface crevice. For the DFT-D optimizations, truncated models of PK083 (*N*-ethylpyrrole) and the p53-Y220C pocket (as depicted) were used (only non-hydrogen atoms and polar protons are shown). Interaction energies of each ligand model were compared to the *N*-ethyl reference interaction energy ΔE to calculate relative interaction energies ($\Delta\Delta E = \Delta E_{Ligand} - \Delta E_{N\text{-ethyl}}$). The three distinct orientations of local minima of the 2-fluoroethyl anchor showed different interactions energies, indicating that orientation of the C-F vector towards the backbone carbonyl groups of Leu145 and Trp146 yields the most favorable interaction energy.

Distances between the backbone amides of Leu145 and Trp146 and the nearest C-F groups were between 3.0 Å and 3.3 Å in the optimized structures, which is in good agreement with typical CF…C=O distances (3.0 Å – 3.7 Å) for orthogonal multipolar interactions.[408] C-F…S and C-F…HS distances ranged from 2.8 Å to 3.6 Å, which is also in good agreement with experimentally observed distances for fluorine-sulfur contacts

(2.8 Å – 3.4 Å) in protein structures.[323] The DFT-D3 calculations indicated that the relative interaction energies (ΔΔE) for all fluorinated ethyl groups improved by at least -2 kcal/mol compared with the *N*-ethyl reference energy ΔE (ΔE = $E_{complex}$ - ($E_{receptor}$ + E_{ligand})) (Figure 3.14B). The energetically most favorable conformation (ΔΔE = -6.5 kcal/mol) of the 2-fluoroethyl group was orientation 1 (Figure 3.14B), in which the C-F vector points towards the backbone amides of Leu145 and Trp146, predicting two potential orthogonal multipolar interactions between the fluorine atom and both carbonyl groups. Orientations 2 and 3 of the 2-fluoroethyl group, where the fluorines were oriented towards the sulfhydryl group of Cys220, were energetically less favorable, with respective ΔΔE values of -2.2 kcal/mol and -3.7 kcal/mol. The relative interaction energy of conformation 1 of the difluoro ethyl moiety (ΔΔE = -6.9 kcal/mol) was similar to the most favored 2-fluoroethyl conformation, whereas the trifluoro-substituted ethyl anchor was energetically less favorable with a ΔΔE value of -4.2 kcal/mol (Figure 3.14B). However, the calculated DFT-D3 energies only yield an estimate of the ligand-protein interaction at the chosen computational level in a model system of small size and neglect other contributions to the free energy of binding such as entropic changes and desolvation penalties.

These results confirm the hypothesis that the cys220 mutational hotspot is potentially fluorophilic, as such, PK083 analogues bearing fluoroethylated groups at the central nitrogen of carbazole (position R, Figure 3.13, Figure 3.14, Figure 3.15) were targeted for synthesis toward a PK083 analogue that shows improved affinity potentially through direct fluorine-protein interactions. Additionally, by variation of the substitution pattern at the level of the pendant amine at R^2, additional information on the effect on the overall potency can be established.

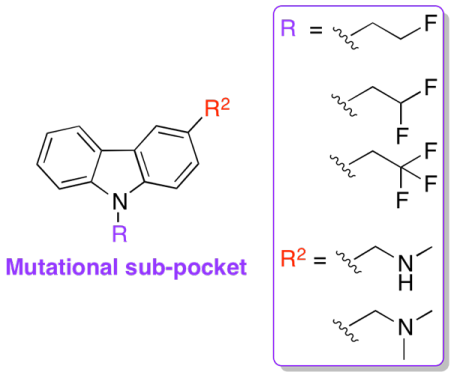


Figure 3.15: Compounds for targeting Cys220 sub-pocket fluorophillicity.

When R = CH_2CH_2F and R^2 = CH_2NHMe, the compound was commercially available and hence, was not synthesized.

3.3 Results and Discussion

The original strategy towards the synthesis of fluoroethylated analogues of PK083 was to use the same synthesis established in Chapter 2.0, starting from 3-bromocarbazole (Figure 3.16). Lithium-halogen exchange with 3-bromocarbazole to generate a lithiated intermediate, followed by reaction with DMF and subsequent hydrolysis to yield carbazole-3-carbaldehyde. Reductive amination using sodium triacetoxyborohydride and dimethylamine yields privileged intermediate **2.37**, which, following a nucleophilic substitution reaction between the generated nitrogen anion and iodo-fluoroethane electrophiles yields final products.

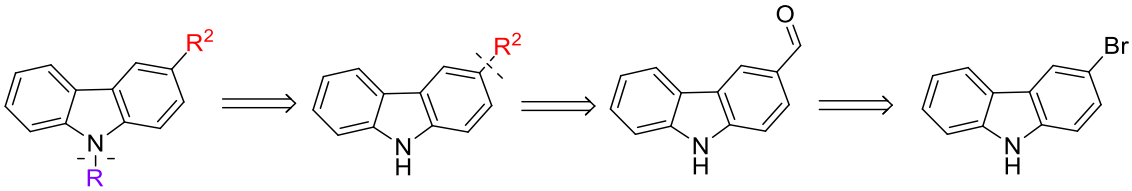
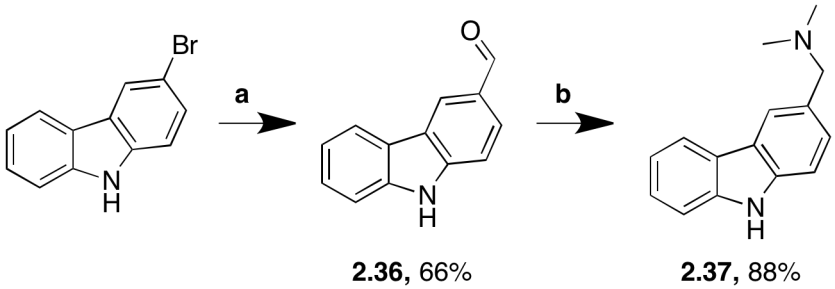


Figure 3.16: Retrosynthetic disconnection towards compounds targeting subsite I and the fluorophilicity of the mutational sub-pocket.

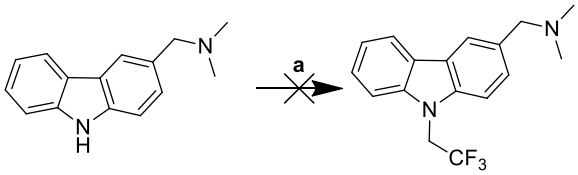
3.3.1 Alkylation Chemistry

The synthesis of intermediate **2.37** was conducted according the conditions established in Chapter 2 (Scheme 3.1).



Scheme 3.1: Synthesis of intermediate **2.37** from 3-bromocarbazole. (a) NaH, THF, rt then -78 °C, *n*BuLi, -78 - -35 °C then -78 °C, DMF, THF, rt; (b) $HNMe_2$, $NaBH(OAc)_3$, THF, rt.

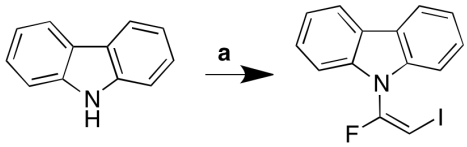
Initial attempts at trifluoroethylation via a S_N2 reaction between the nitrogen nucleophile, and 1-(9-carbazol-3-yl)-*N,N*-dimethylmethanamine (**2.37**) led to an inseparable mixture (Scheme 3.2).



2.37 **3.7a**

Scheme 3.2: Attempted synthesis of final product **3.7a** from intermediate **2.37**. (a) NaH, DMF, rt, 30 min then ICH_2CF_3, rt, 18 h.

This alkylation reaction was further investigated by examining the analogous reaction of carbazole with 1-iodo-2,2,2-triflurorethane using sodium hydride as a base at room temperature. This yielded the unusual vinylic elimination product **3.8** (Scheme 3.3) the structure of which was determined by ^{1}H NMR, HRMS-ESI and X-ray crystallography. The experiment was repeated three times under identical conditions and the outcome of the reaction was identical. Similar observations of vinylic elimination products formed upon trifluoroethylation have been made by Suehiro *et al*. in the radiosynthesis of a tumour hypoxia marker.[447] By varying the base, it was found that only the completion of the reaction was affected (Table 3.13). It is of note that the yield of the reaction towards the elimination product when using a 'hard' base was significantly higher than when using a 'soft' base such as cesium carbonate. Thus, it cannot be discounted that the nature of the base may affect the outcome of the reaction. This reaction proceeded with a high degree of isomeric control, with exclusive formation of the E isomer (as determined by ^{1}H NMR and X-ray crystallography).



3.8, 40%

Scheme 3.3: Elimination reaction. (a) NaH, DMF, rt then ICH_2CF_3, 18 h.

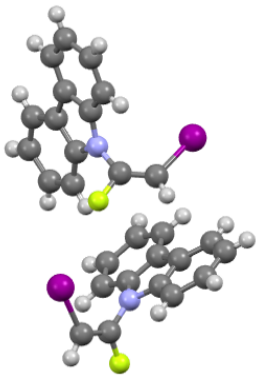


Figure 3.17: Crystal structure of vinylic elimination product **3**.8 as determined by X-ray crystallography.

	Base	Solvent	Conditions	Yield		
				3.8	Carbazole	Trifluoroethylated product
A	NaH	DMF	Room temperature	40	60	0
B	Cs_2CO_3	DMF	Room temperature	0	100	0
C	PS-BEMP	DMF	Room temperature	0	100	0
D	Et_3N	DMF	Room temperature	0	100	0
E	NaH	DMF	150°C, microwave irradiation, 30 minutes	69	27	4
F	Cs_2CO_3	DMF	150°, microwave irradiation, 30 minutes	12	81	6
G	Et_3N	DMF	150°, microwave irradiation, 30 minutes	0	100	0

Table 3.13: Reaction optimisation of trifluoroethylation of carbazole. Yields determined by ^{1}H NMR (DMSO-d_6). The trifluoroethylation reagent was 1-iodo-2,2-trifluoroethane in each case.

Using microwave-assisted chemistry it was found that the trifluoroethylated product could be isolated, ableit in a low yield (Table 3.13, entries E-G). This implies that the elimination reaction has a lower enthalpy of reaction than for the simple S_N2 reaction. This is contrary to initial expectations as the enthalpy of breaking two C-F bonds (880 kJ mol^{-1}) is far in excess of one C-I bond (240 kJ mol^{-1}). It has been shown by fluorination studies, that the effect of replacing each of the C-H bonds on C2 of iodoethane with C-F bonds effectively decreases the electrophilic character of C1 as the electrophilic character of C2 increases exponentially.

The competition between nucleophilic attack at C1 and C2 was circumvented in favour of nucleophilic attack at C1 by the use of trifluoroethyltosylate.[447] As the tosylate is the 'harder' electrophile it is more reactive in a S_N2 displacement and less reactive in E2 compared with the iodide and hence nucleophilic attack took place solely at C1 (Figure 3.18).

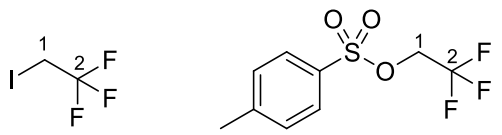
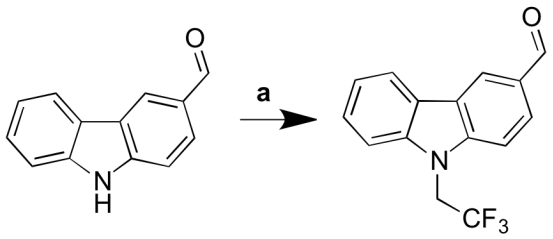


Figure 3.18: Structures of trifluoroethylation reagents used.

The optimum conditions were obtained when using 9*H*-carbazole-3-carbaldehyde as a starting material (electron-withdrawing aldehyde at 3- position) with Cs_2CO_3 or NaH as

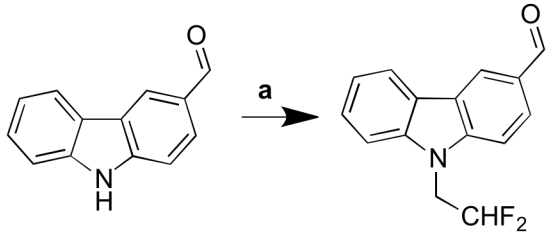
a base, DMF as a solvent and stirring under microwave irradiation at 150°C for 30 minutes (Scheme 3.4).



2.36 **3.9a,** 50%

Scheme 3.4: Reaction scheme for the optimised procedure for trifluoroethylation. (a) $TsOCH_2CF_3$, Cs_2CO_3, DMF, microwave irradiation, 150 °C, 30 min.

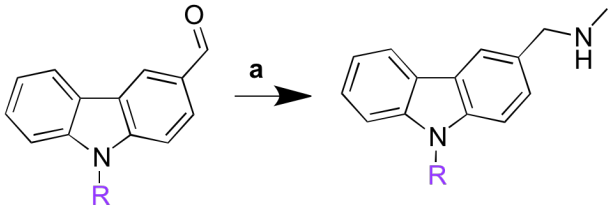
These optimum conditions were extended to N-alkylation with 1-iodo-2,2-difluoroethane as a reagent, which interestingly did not show the same reverse reactivity as for 1-iodo-2,2,2-trifluoroethane (Scheme 3.5).



2.36 **3.9b,** 44%

Scheme 3.5: The generic reaction scheme for the synthesis of mono- and di-fluoroethylated aldehydes.

Analogues of PK083, where $R^2 = CH_2NHMe$ were synthesised in parallel by reductive amination from the respective aldehydes to give the final products **3.7a** and **3.7b** (Scheme 3.6).



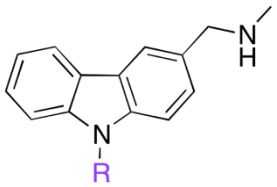
3.9a, R = CH_2CF_3 **3.7b**, R = CH_2CF_3, 40%
3.9b, R = CH_2CHF_2 **3.7c**, R = CH_2CHF_2, 58%

Scheme 3.6: General procedure for the reductive amination of N-fluorinated aldehydes **3.9a** and **3.9b** to give final products **3.7b** and **3.7c**.(a) $Ti(O^iPr)_4$, $MeNH_2$, NEt_3, DCM, EtOH 1:1, 18 h, rt then $NaBH_4$, 8 h , rt.

3.3.2 Effect of the substitution at R on Potency

The stabilizing effect of compounds **3.7a** and **3.7b** on p53-Y220C, along with the mono-fluorinated analogue that was commercially available, were tested by DSF and ITC and

are summarized in Table 3.14.[446] Correlation between the two assays was adaquate, showing a negative correlation of -0.75.



Compound	R	DSF ΔT_m [125 µM]	ITC K_d (µM)
PK083	CH_2CH_3	0.8	124
PK9241	CH_2CH_2F	0.4	138 ± 35.2
PK9254 (**3.7c**)	CH_2CHF_2	0.3	101 ± 10.8
PK9255 (**3.7b**)	CH_2CF_3	1.2	28 ± 2.6

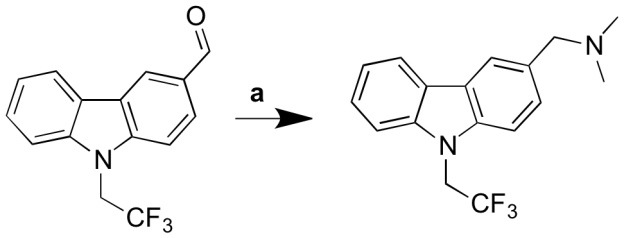
Table 3.14: DSF thermal shift and ITC K_d determination for PK083, and N-fluoroethylated analogues PK9241, PK9254 (**3.7c**) and PK9255 (**3.7b**). DSF and ITC assay conditions are described in Chapter 6. Correlation between DSF stabilization and ITC dissociation constant = -0.75.

The N-2,2,2-trifluoroethyl substituted carbazoles PK9255 (**3.7b**) increased stability of the p53-Y220C DBD by 1.2 K at a compound concentration of 125 µM, showing a clear improvement over their N-ethyl substituted counterparts PK083. In contrast, the difluorinated (**3.7c**) and mono- fluorinated carbazoles PK9254 and PK9241 were significantly worse than the parent compound in terms of protein stabilization, with respective thermal shifts of 0.3 and 0.4 K.

The trifluoro-substituted carbazole PK9255 (**3.7b**) was the most potent compound and showed ITC KD value of 28 µM (Table 3.14), which corresponds to a 5-fold increase in affinity compared to PK083. Despite prediction of improved interaction energies by DFT-D calculations, the mono- and difluorinated analogues, PK9241 and PK9254 (**3.7c**), failed to display improved affinities. They had K_d values of 101 µM and 138 µM, respectively.

3.3.3 Variation at R^2

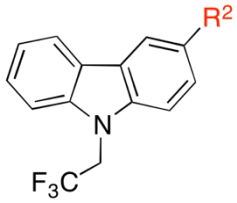
A dimethylated benzylic amine analogue of the most potent trifluoroethyl substituted carbazole, PK9255 (**3.7b**) was synthesized from intermediate **3.9a** by reductive amination with dimethylamine (Scheme 3.7).



3.9a **3.7a,** 56%

Scheme 3.7: Reductive amination of **3.9a** with dimethylamine to yield the PK9255 analogue **3.7a**. (a) $Ti(O^iPr)_4$, $HNMe_2$, NEt_3, DCM:EtOH 1:1, 18 h, rt then $NaBH_4$, rt, 8 h.

The stabilizing effect of **3.7a** on p53-Y220C was examined by DSF and the dissociation constant was determined by ITC and are summarized in Table 3.15.



Compound	R^2	DSF ΔT_m [125 µM]	ITC K_d (µM)
PK9255 (**3.7b**)	CH_2NHMe	1.2	28 ± 2.6
MB065 (**3.7a**)	CH_2NMe_2	1.2	37

Table 3.15: DSF thermal shift and ITC K_d determination for N-fluoroethylated analogues PK9255 (**3.7b**) and MB065 (**3.7a**). DSF and ITC assay conditions are described in Chapter 6.

The substitution pattern of the solvent-exposed amine had a minor effect on affinity, with the additional methyl group in 5 resulting in a 30% drop in affinity.

3.3.4 X-Ray crystallography

The binding mode of the carbazoles was almost identical to that of PK083, however, variations at R^2 caused the benzylic amine to flip, breaking the hydrogen bond with the backbone Asp228 C=O, forming a new hydrogen bond with the C=O of Val147 (Figure 3.19E). Hence, small differences in the orientation of the central carbazole scaffold can be attributed to the different methylation state of the amine moiety. Upon monofluorination, the fluorine atom points toward the carbonyl groups of Leu145 and Trp146 with a F··· C=O interaction distance of 3.4 Å (Figure 3.19A). The C−F vector and the planes of the backbone carbonyl group of Leu145 and Trp146 show a nearly orthogonal arrangement, with respective O=C···F angles of 97.5° and 80.1°, which is in good agreement with O=C···F angles found in the PDB and CSD for orthogonal multipolar interactions. This preferred orientation of the C−F vector in the crystal structure is also in good agreement with initial DFT-D predictions (Figure 3.14). Similar to the Y220C-PK083 complex, the side chain of Cys220, in immediate vicinity of the ethyl anchor, adopts two alternative, albeit very similar conformations. For the N-2-

difluoroethyl anchor, there are two alternative conformations in the p53-Y220C binding pocket (Figure 3.19B, C). In both conformations, one fluorine atom interacts with the

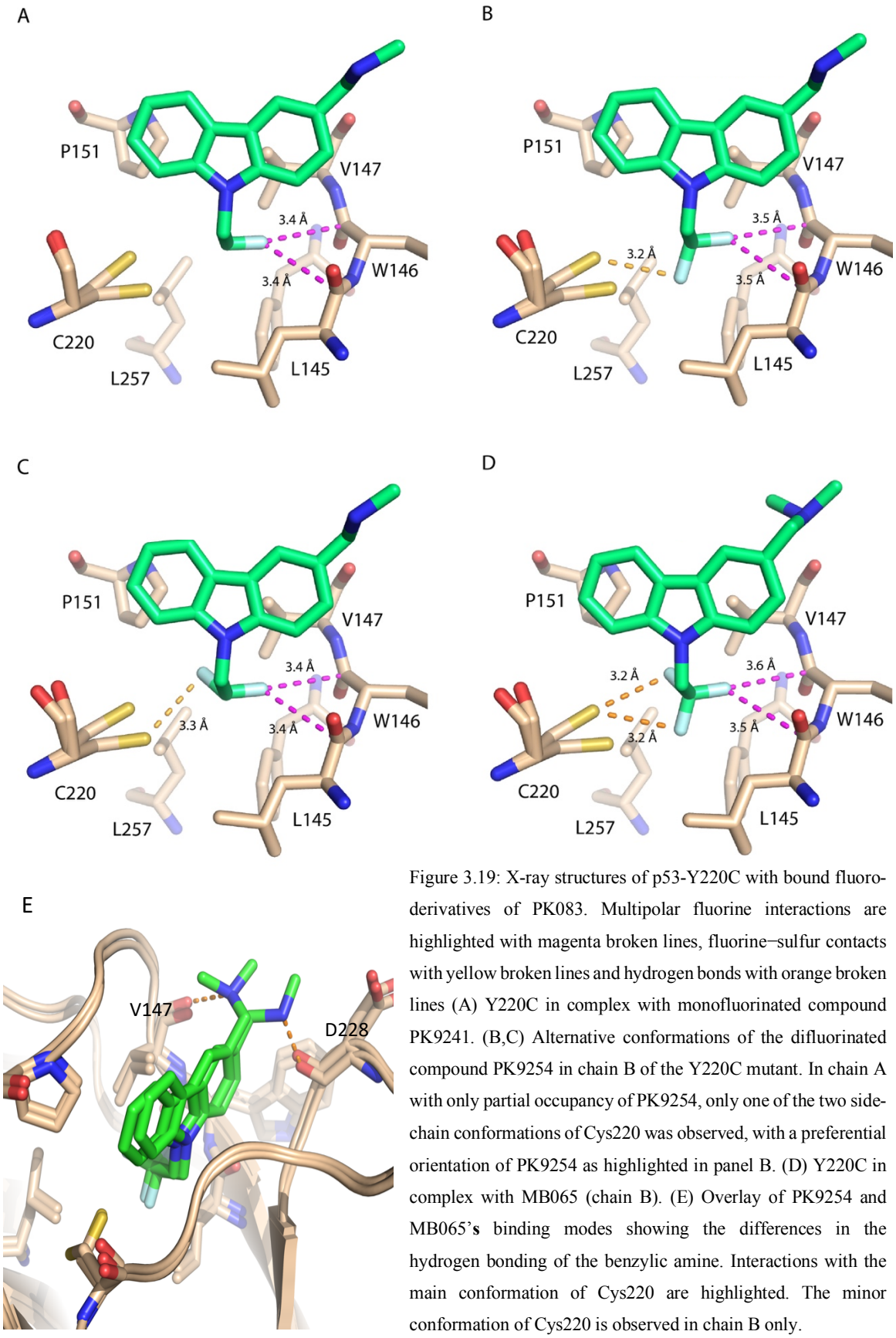


Figure 3.19: X-ray structures of p53-Y220C with bound fluoro-derivatives of PK083. Multipolar fluorine interactions are highlighted with magenta broken lines, fluorine–sulfur contacts with yellow broken lines and hydrogen bonds with orange broken lines (A) Y220C in complex with monofluorinated compound PK9241. (B,C) Alternative conformations of the difluorinated compound PK9254 in chain B of the Y220C mutant. In chain A with only partial occupancy of PK9254, only one of the two side-chain conformations of Cys220 was observed, with a preferential orientation of PK9254 as highlighted in panel B. (D) Y220C in complex with MB065 (chain B). (E) Overlay of PK9254 and MB065's binding modes showing the differences in the hydrogen bonding of the benzylic amine. Interactions with the main conformation of Cys220 are highlighted. The minor conformation of Cys220 is observed in chain B only.

carbonyl groups of Leu145 and Trp146 in an almost identical fashion as observed for the monofluorinated analogue. The second fluorine interacts with the thiol group of Cys220, pointing either toward Pro151 (Figure 3.19C) or toward Leu145 at the bottom of the binding pocket (Figure 3.19B), which is essentially the result of a 120° rotation around the ethyl anchor C−C bond. Interestingly, only one of the two Cys conformations was observed in chain A, coinciding with a preferential orientation of the difluoroethyl anchor in the orientation highlighted in Figure 3.19, whereby the fluorine forms a weak hydrogen bond with the SH group of Cys220. In both orientations, the fluorine atom interacts with the thiol group at a distance of about 3.2 Å, which is more or less the sum of the van der Waals radii of sulfur and fluorine (rF = 1.47 Å; rS = 1.80 Å). In the structure of the most potent trifluoromethylated compound, the CF_3 group aligns well with the different fluorine positions observed for monofluorinated and difluorinated (Figure 3.19D); i.e., it interacts with the backbone carbonyl groups of Leu145 and Trp146 as well as with the thiol group of Cys220. Analysis of C−S···F angles in the complexes with di- and trifluorinated compounds suggests that fluorine interacts with Cys220 via weak hydrogen bonding with the polarized proton of the thiol function and via sulfur σ-hole bonding at an angle close to 180 °.

3.4 Conclusions

Incorporation of fluorine atoms into the *N*-ethyl anchor of the p53-Y220C stabilizer PK083 yielded two compounds, PK9255 and MB065, with substantially increased p53-Y220C stabilization (1.2 K at 125 μM) and Y220C binding affinity. Compared with the parent compounds PK083 and PK8017, the ITC K_D values improved by a factor of 5 and the free energy of binding by approximately -0.9 kcal/mol (at 293K). Typically, C-F…C=O orthogonal multipolar interactions with optimal geometry raise the binding free enthalpy by about -0.2 to -0.3 kcal/mol,[324] which suggests that the observed C-F…C=O interactions with the carbonyl groups of Leu145 and Trp146 are the main factor for the potency increase of the trifluorinated PK083 derivatives. The DFT-D calculations are consistent with the binding modes observed in the crystal structures, as the preferred orientation of fluorine towards the backbone carbonyl groups of Leu145 and Trp146 in these structures indicates that interactions of the protein carbonyl groups with fluorine atoms contribute more strongly to the affinity increase than interactions with the thiol group of Cys220 or apolar protons of hydrophobic side chains. The preferred interaction with the protein backbone then seems to direct the orientation of additional fluorine atoms

in the binding pocket, interacting with the thiol group of Cys220 via either a hydrogen bond or a planar sulfur σ-hole interaction.

Although DFT-D calculations predicted improved interaction energies for mono-fluorinated and di-fluorinated PK083 derivatives, this was not reflected in the experimentally determined dissociation constants and stability data for compounds PK9241 and PK9254, which were close to that of the parent compound PK083. These discrepancies between theoretical and experimental data are likely due to different desolvation penalties associated with each fluorinated group. Fluorine can act as a weak hydrogen-bond acceptor in a 2,2-difluoromethyl group and even more strongly in a 2-fluoromethyl group, although hydrogen-bond strength was found to be significantly weaker than for the conventional hydrogen-bond acceptor acetophenone.[433] For example, difluoroalkyl groups have been used as thiol surrogates in drug discovery because of their similar steric properties and the acidity of the terminal hydrogen resulting from the high polarization of the C-H bond by the geminal fluorine atoms.

In conclusion, targeting Cys220 and the backbone carbonyl groups of Leu145 and Trp146 via fluorine bonding helped to further optimize the carbazole scaffold. *N*-3,3,3-Trifluoroethyl substituted carbazoles PK9255 and MB065 exhibit a high ligand efficiency (LE = 0.3 kcal mol^{-1} $atom^{-1}$), although binding is still relatively weak for these fragment-like molecules. Differences in the methylation state of the benzylic amine had a minimal effect on the overall potency, where dimethylation resulted in a reduction in affinity, measured by ITC dissociation constant of 37 μM vs 28 μM. Crystallography revealed that dimethylation resulted in flip of the benzylic amine, breaking the hydrogen bond observed in PK083 analogues between the amine and the backbone C=O of Asp228, forming a new hydrogen bond between the protonated amine (HCl) and the backbone C=O of Val147. This provides a rational of the difference in affinity between dimethylated analogues (e.g. PK8017) and monomethylated PK083.

Chapter 4 :Targeting Subsite II

Efforts towards improving PK083 potency towards p53-Y220C via targeting the mutant specific sub-pocket for affinity improvements resulted in the synthesis of PK9255, which shows enhanced affinity through direct fluorine-protein interactions resulting in a K_d = 28 μM as determined by ITC.[446] Subsequent efforts to further enhance the selectivity and potency of PK083 through the design of analogues that bear covalent modification warheads were unsuccessful (Dr Adam Close, unpublished results). Previous attempts to target the solvent exposed subsite I through functionalization at R^2 did not exhibit any patterns within the SAR. Furthermore, initial attempts to target the transiently open subsite III were unsuccessful. Hence, alternative areas of the Y220C pocket were examined for areas where potential affinity gains could be made. By examination of the x-ray structure of PK083 bound to the Y220C pocket, the position R^3 was identified for growth that allows targeting the proline-rich subsite II (Figure 4.1).[261]

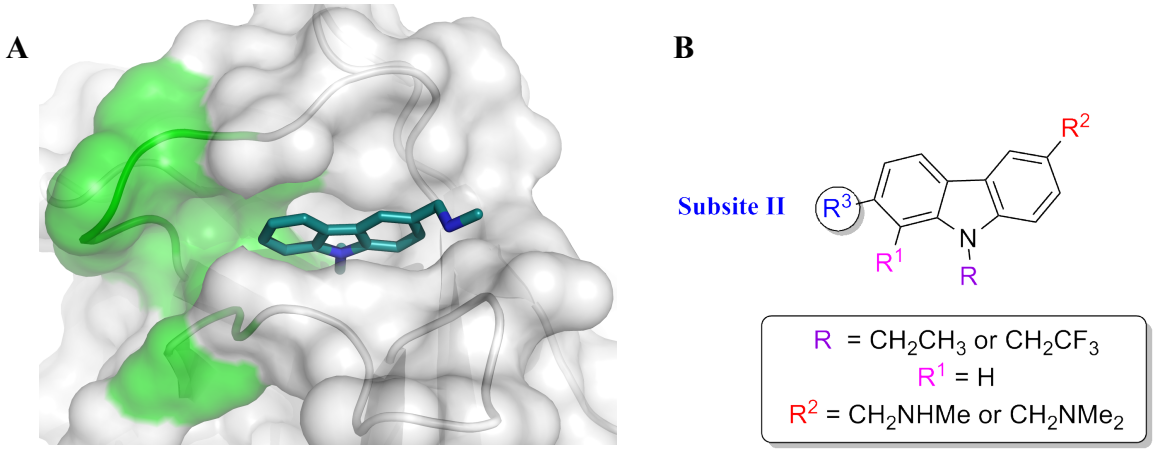


Figure 4.1: Summary of the strategies towards improving the affinity of PK083 analogues towards p53-Y220C. (A) X-ray structure of PK083 bound to p53-Y220C (PDB: 2VUK) showing subsite II (highlighted in green). (B) Generic chemical structure of compounds synthesised to target subsite II.

This chapter describes the synthesis of compounds that target the proline rich subsite II subpocket through the functionalization of R^3, and, using SAR and x-ray crystallography, the optimization towards higher affinity PK083 analogues, where the most potent substitutions occurred when R^3 is a 5-membered heterocycle (Figure 4.1B).

4.1 Properties of Subsite II

Subsite II is accessed via bridging a narrow gap from the central cavity, formed by Cys220 and Pro151. This sub-pocket is composed of a hydrophobic interaction surface

formed by Pro151, Pro152 and Pro153 as well as the backbone carbonyl of Cys220. This hydrophobic surface is neighboured by two polar neutral residues, Thr155 and Glu221. A structural water (W1) occupies the pocket sitting between Cys220 and Thr155 (Figure 4.2).[261]

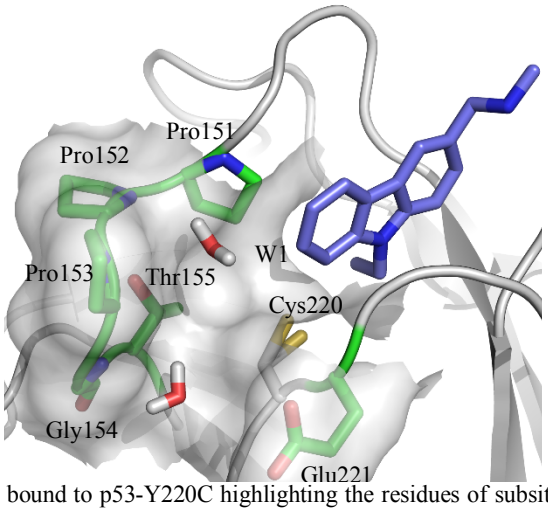


Figure 4.2: Structure of PK083 bound to p53-Y220C highlighting the residues of subsite II, carbons shown in green (subsite II).

Computational studies, such as MD simulations and wrapping studies, have sought to probe subsite II, seeking to identify the most efficient means of targeting it.[291,448] The results of these studies are summarized below.

4.1.1 MD Simulations

The use of molecular simulations to probe the dynamic properties of proteins using isopropanol solvation studies has been demonstrated to be a useful tool in determining drug-binding sites and protein-protein interaction interfaces on proteins of pharmacological interest.[449] For example, the binding mode of two peptide substrates of BCL-6. Isopropanol is chosen as it contains both polar and non-polar properties that are intended to simulate hydrogen bonding and hydrophobic interactions similarly to that of a ligand.[450] Basse and Kaar *et al*. utilized this technique to probe the potential ligand binding 'hotspots' of p53-Y220C.[291]

Simulations were performed in a water isopropanol mixture using an isopropanol concentration of 20% v/v to assess preferential solvation of the protein. Analysis of the isopropanol density isosurface of p53-Y220C found several regions of high density, the highest of which was found on the DNA binding domain and a surface patch around

Leu188 and Leu201.[451] However, the fragment screen of small molecules discussed in Section 2.3 identified that small molecules tend to bind only the mutation site, hence the simulation was restricted to only the Y220C pocket.[291]

Isopropanol density in the mutation induced cavity was over four times higher than the expected value for the bulk mixture, indicating the propensity of solvent binding, and hence ligand binding in the pocket. Localization of isopropanol molecules within the cavity occurred between Pro153, Pro220 and Cys220 was observed. The hydroxyl group of isopropanol formed hydrogen bonds with the backbone carbonyl of Pro151 and Pro152 with an estimated free energy of binding of -3.9 kcal/mol. Conversely, the area of the cavity between side chains of Val147, Pro151 and Pro223, the area usually occupied by aromatic rings of fragments that bind Y220C, was devoid of isopropanol (Figure 4.3). The authors reasoned that this was due to the increased steric bulk of isopropanol compared with planar aromatic rings. These identify subsite II as an area that shows druggability.

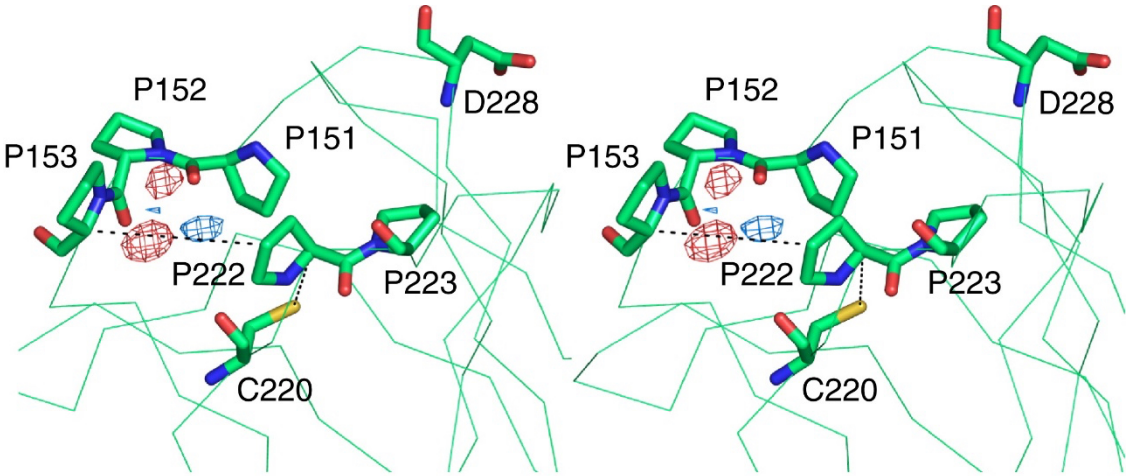


Figure 4.3: The density of isopropanol in the cavity contoured at 4 times the density of the bulk mixture. Isopropanol oxygen is shown in red and carbon is shown in blue. The distances used to monitor the structural motions are shown in dashed black lines. Reproduced from Chemistry and Biology, N. Basse, J. L. Kaar, G. Settanni, A. C. Joerger, T. J. Rutherford, and A. R. Fersht, 'Toward the Rational Design of p-53-Stabilizing Drugs: Probing the Surface of the Oncogenic Y220C Mutant,' 46–56, Copyright (2010), with permission from Elsevier.

As well as identifying potential binding sites, molecular dynamics simulations allow the examination of how structural motions in the cavity affect ligand binding. Analysis of the cavity dynamics found that the distance between C_{α}Pro153 and C_{α}Pro222 varied between 7-9 Å over the time course of the simulation. In the absence of isopropanol, the cavity fluctuated between 5-7 Å, corresponding to an apparent open and closed state.

Additionally, the depth of the cavity, which was taken as the distance between the sulfur of Cys220 and C_αPro222, fluctuated in a range comparable with crystal structures.[291] However, the depth was only significantly altered when isopropanol occupied the cavity indicating an induced-fit phenomenon. These results indicate that the central cavity and subsite II are highly dynamic.

Furthermore, Basse and Kaar used Gaussian Network Model (GNM) analysis to assess whether the binding of fragments alters the plasticity of the cavity.[291] GNM represents the protein as an elastic network where harmonic springs connect neighbouring nodes centred on the C_α atoms of residues. GNM has been proven to be highly accurate in reproducing experimental data on structural fluctuations in biological macromolecules and is widely used to study the functional consequence of such fluctuations.[452–455] The results of the GNM analysis indicated that the presence of fragments markedly reduce the flexibility of the pocket. The backbone dynamics of p53-Y220C are virtually identical to the rest of the protein, thus enhancing the rigidity of the region. The dynamics in the absence of the fragments are more flexible. The authors reasoned that this could be attributed to the fragments coupling the proline-rich loops that line to mutational cavity to the rest of the protein, thus enhancing the rigidity. This is supported by the observation that Pro222 exhibits significantly reduced temperature factors in the crystal structures of all Y220-ligand complexes compared to structures of the ligand free mutant. Thus, targeting subsite II with hydrophobic groups seeking to optimally occupy the sub-pocket, should reduce flexibility in the proline loop, enhancing rigidity, reducing entropy and improving the temperature sensitive phenotype that this thermodynamically destabilizing mutation causes.

4.1.2 Wrapping Effects in the Y220C Pocket

The wrapping concept was introduced by Prof. Ariel Fernandez, and states that to prevail in water environments, soluble proteins must protect their backbone hydrogen bonds from the disruptive effect of water attack by clustering non-polar residues around them.[456–464] Underwrapped interactions, termed dehydrons, represent thermodynamically unstable, vulnerable sites where the level of underwrapping has been found to correlate with the degree of structural disorder.[456,461,464] Dehydrons are adhesive, hence, are promoters of molecular associations because their inherent stability increases upon approach of

additional non-polar groups. Consequently, dehydrons constitute key motifs that signal protein binding sites.

Accordino *et al*. studied the impact of the Y220C mutation in terms of wrapping.[465] In the wild-type structure, a wrapping interaction is formed between Y220 and Thr155. The tyrosine to cysteine mutation results in the loosening of the protein structure, allowing water to access the dehydronic Cys220···Thr155 hydrogen bond formed between the backbone NH of Cys220 and the side-chain hydroxyl of Thr155 that now wildly fluctuates, breaking and reforming over the course of time, thus promoting structural destabilization (Figure 4.4A). PK083 only provides partial wrapping as it does not extend into subsite II, leaving the hydrogen bond vulnerable to attack. Computational docking experiments of PK083 analogues that bear groups that extend into subsite II exhibited an increased free energy of binding compared with PK083 (Figure 4.4B).[465]

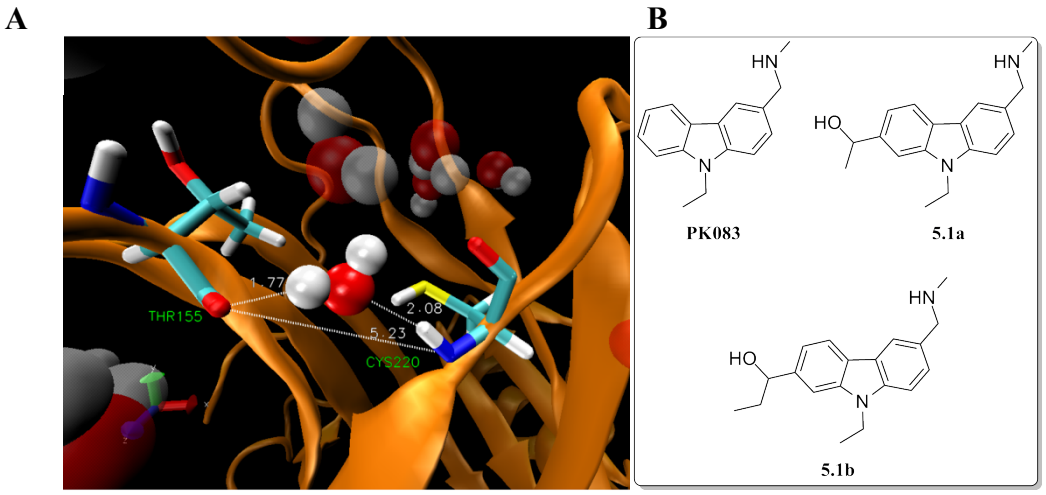


Figure 4.4: Figure 4: (A) Structure of p53-Y220C (PDB: 2J1X) showing the dehydronic Cys220-Thr155 hydrogen bond being disrupted by water.[448] Reproduced from Ref. 448. (B) Chemical structure of computationally docked PK083 analogues.

4.1.3 Summary of Computational Results

The computational results suggest that tyrosine to cysteine mutation results in the loss of the stabilizing influence of the tyrosine on the surrounding proline loops, i.e.subsite II, leading to much greater flexibility, as suggested by MD simulations and the GNM analysis. Additionally, the mutation results in the Cys220-Thr150 dehydronic interaction, which is vulnerable to attack from water. Thus, the mutation seems to destabilize subsite II, undoubtedly contributing to the thermodynamically unstable phenotype that p53-Y220C exhibits (greater flexibility of proteins is linked to instability).[466] These results

suggest targeting subsite II, by enhancing rigidity and providing compensation for the dehydronic Thr155-Cys220 interaction, is a viable method of improving PK083 affinity and potency.

Initial docking results suggest that significant affinity gains can be made by targeting subsite II.[465] Below, a summary of fragments that bind Y220C, occupying subsite II are discussed, and SAR findings are derived.

4.2 Compounds that bind subsite II

The available SAR around compounds that target subsite II is substantially smaller than that for the central cavity and subsite I. The compounds that bind subsite II can be classified broadly as (i) fragments that were discovered as part of a high throughput screen, and (ii) iodophenols that were rationally designed to target subsite II.[179]

4.2.1 Fragments

As previously discussed in Section 2.3, a fragment screen of 1900 chemically diverse compounds against T-p53-Y220C, using water LOGSY, thermal denaturation scanning fluorimetry and $^{15}N/^{1}H$ HSQC, identified 252 hits, 87 of which were confirmed.[291] Of the latter, several aromatic fragments, upon x-ray crystallography studies, were found to bind to both the central cavity and subsite II (Figure 4.5A).[258] However, in most cases, the electron density for the fragment occupying subsite II was lower, indicating partial occupancy and lower affinity for subsite II.

For example, fragment **4.1** binds deep in the central cavity, with its hydroxyl group forming hydrogen bonds with the backbone C=O of Leu145 as well as a conserved water molecule that bridges Asp228 and Val147. The chlorine faces the narrow hydrophobic channel that leads to subsite II. Additional hydrogen bonds are formed between the amide C=O and the side chain hydroxyl group of Thr150 as well as the amides NH and the structural water. The piperidine moiety points to solvent. The second molecule sits in subsite II, where its aromatic ring packs against Pro153 and Pro222. The chloro moiety points towards Cys220 (Figure 4.5C).[258]

Similar binding is observed for fragment **1.1**, in which the fragment occupying the central cavity sits deep in the central cavity between Pro223 and Val147, where the ethyl group is buried deep in the hydrophobic mutation site. A hydrogen bond is formed between the indole NH and the hydroxyl group of Leu 145. Similarly to **4.1**, the piperidine moiety of **1.1** points direct to solvent. The second fragment that occupies subsite II is also engaged in hydrophobic interactions with Pro153 and Pro222, with the aromatic core of indole packing against them. The NH of indole forms a hydrogen bond with the backbone carbonyl of Pro152 (Figure 4.5B).[258]

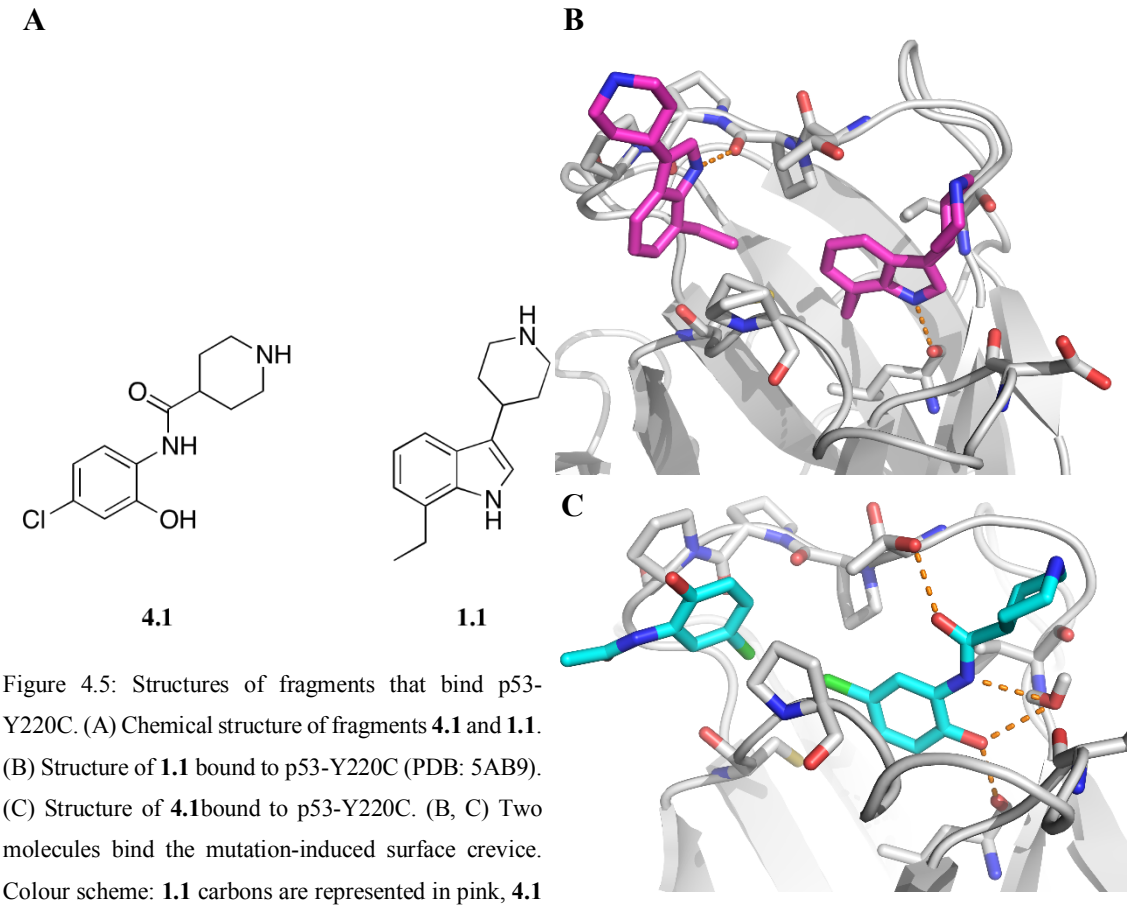


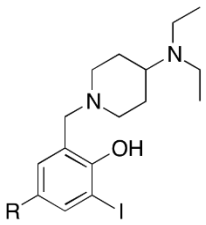
Figure 4.5: Structures of fragments that bind p53-Y220C. (A) Chemical structure of fragments **4.1** and **1.1**. (B) Structure of **1.1** bound to p53-Y220C (PDB: 5AB9). (C) Structure of **4.1**bound to p53-Y220C. (B, C) Two molecules bind the mutation-induced surface crevice. Colour scheme: **1.1** carbons are represented in pink, **4.1** carbons are represented in cyan. Hydrogen bonds are shown as dashed orange lines.

4.2.2 Iodophenols

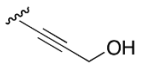
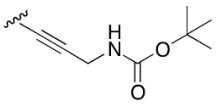
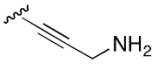
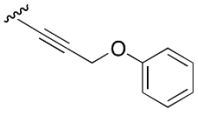
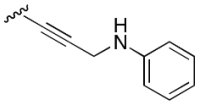
The iodophenol (hit compound, **1.2**) series was discovered from the same fragment screen, where the parent compound does not bind subsite II.[291] However, in the process of hit-to-lead optimization, rational design of compounds that grow the fragment into subsite II resulted in marked improvements in affinity (Figure 4.6A).[179] The fragment

was grown via an acetylene linker, chosen as it possesses the necessary rigidity and planarity to bridge the narrow channel formed by Cys220 and Pro151 into subsite II and is synthetically easy to install via a Sonogashira coupling reaction.

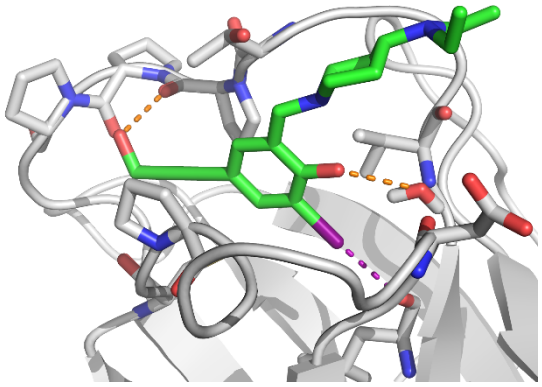
A



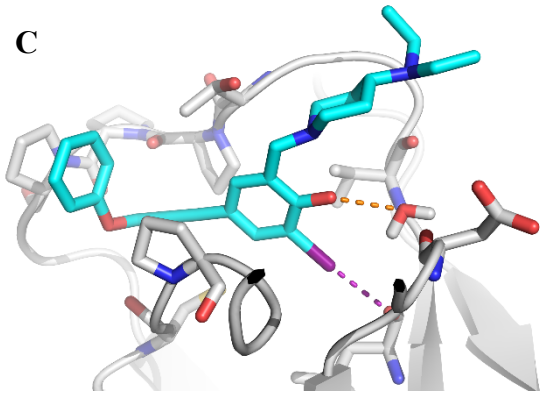
4.2

Compound	R	ITC K_D (μM)
4.2a	-I	105
4.2b		107
4.2c		15.5
4.2d		1080 (NMR)
4.2e		20.6
4.2f		9.7

B



C



D

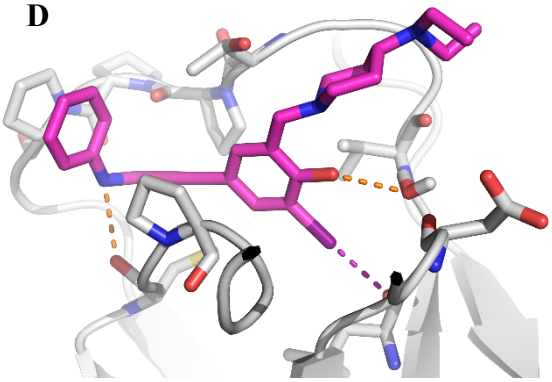


Figure 4.6: SAR of iodophenols that target subsite II via acetylene groups. (A) Chemical structures and SAR of **4.2**. (B) Structure of **4.2b** bound to p53-Y220C (PDB: 4AGN); (C) Structure of **4.2e** bound to p53-Y220C (PDB: 4AGP); (D) Structure of **4.2f** bound to p53-Y220C. Colour scheme: Hydrogen bonds are represented as orange dashed lines, halogen bonds are represented as purple dashed lines.

Compound **4.2b**, which bears a propargylic alcohol, showed similar binding affinity to the parent compound, where the authors reasoned that this could be attributed to the loss of the bulky iodine being compensated by the interactions formed by the new substituent (Figure 4.6B). Surprisingly, comparison of the Boc-protected propargylamine **4.2c** with the free amine **4.2d** reveals that greater affinity gains were achieved through the introduction of the Boc-protected compound. The additional desolvation penalty

associated with the highly polar free-amine account for this observation. Introduction of substituents **4.2e** and **4.2f** that bear aromatic moieties resulted in high improvements in affinity (Figure 4.6A).

Examination of the structures reveals that, similarly to that observed for fragments **1.1** and **4.1**, favourable CH-π stacking of the phenyl rings against Pro153 occurs (Figure 4.6C, D). Substitution of the free phenoxy substituent in **4.2e** for the aniline substituent in **4.2f** resulted in the most potent substitution. Examination of the structure of **4.2f** bound to p53-Y220C reveals that this is due to the formation of a hydrogen bond between the aniline NH and the backbone carbonyl of Pro222 (Figure 4.6D).

4.2.3 Summary of SAR of subsite II binding compounds

From the above results, the compounds that occupy or target subsite II successfully tend to be hydrophobic, most commonly aromatic in nature. Affinity gains tend to be achieved via preferential CH-π stacking against Pro153, as well as picking up on hydrogen bonding interactions with the backbone carbonyls.

4.3 Design of a Key Intermediate for Targeting Subsite II in the PK083 series

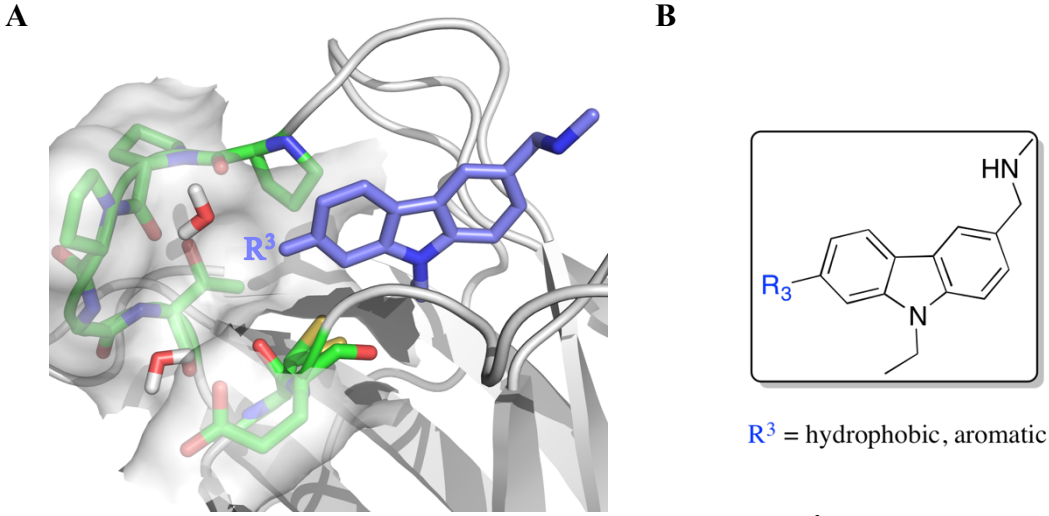
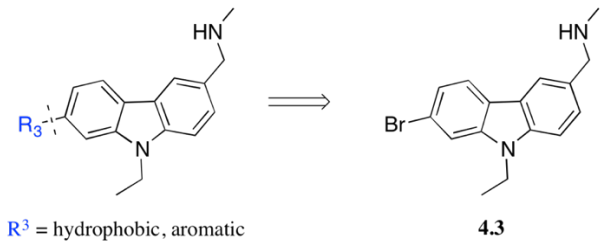


Figure 4.7: (A) Structure of PK083 bound to p53-Y220C (PDB: 2VUK) showing vector R^3 for targeting subsite II. (B) Generic chemical structure of PK083 analogues to target subsite II.

From visual examination of the crystal structure of PK083 bound to p53-Y220C, subsite II can be targeted best through growing PK083 at position R^3, therefore, the required analogues need to possess a 2,6,9 substitution pattern (Figure 4.7A, B).[261] Chemically,

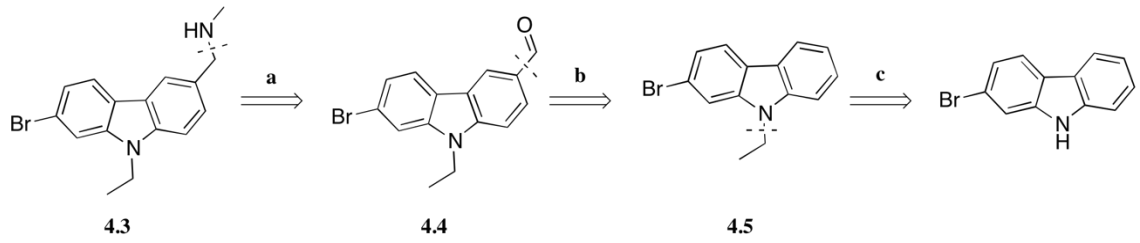
direct functionalization of carbazole to yield a 2-substituted framework is not simple. As discussed in Section 2.5.1, carbazoles are reactive towards electrophiles at positions 1,3,6 and 8.[338] 2-Isomers are rarely obtained. Therefore, the simplest route to 2-substituted carbazoles is via construction of the carbazole using a ring-closing synthesis of the types described in Section 2.5.2, where the starting material for such a reaction bears a substitution pattern that will yield a 2-substituted product.[361,467,468]

Furthermore, using the SAR established in Section 4.2, the retrosynthetic disconnection of the generic structure shown in Figure 4.7B must result in a privileged intermediate that allows for the chemistry to install hydrophobic and aromatic compounds. For example, by using an aryl bromide intermediate in metal-catalysed cross coupling reactions such as the Suzuki Miyaura or Sonogashira reactions, similarly to the approach adapted for the iodophenol series (Scheme 4.1).[469]



Scheme 4.1: Retrosynthetic disconnection of a generic PK083 analogue that shows the required regiochemistry to target subsite II to a privelged intermediate **4.3**.

Synthetically, **4.3** can be synthesised starting from commercially available 2-bromocarbazole according to Scheme 4.2.



Scheme 4.2: Retrosynthesis of **4.3**. (a) $H_2NMe{\cdot}HCl$, Et_3N, $Ti(O^iPr)_4$, DCM/EtOH, rt 18 h then $NaBH_4$, rt, 8 h; (b) $POCl_3$, DMF, Δ; (c) NaH, THF, rt then EtI.

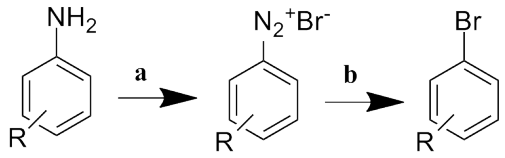
Starting from 2-bromocarbazole, ethylation under similar conditions, decribed in Chapters 2 and 3, yields the intermediate **4.5** that, when subjected to Vilsmeier-Haack formylation conditions, should react in the most electron-rich position, which, due to the inductive effect of the 2-bromo substitution, should give the 6- isomer as the major

product.[470] Reductive amination of **4.4** yields the privileged intermediate **4.3** that can be used in a variety of reactions to yield a sterically and electronically diverse selection of compounds to probe subsite II. Below, a short review of these reactions is presented, concentrating on the Suzuki-Miyaura reaction.

4.4 Reactivity of Aromatic Bromides

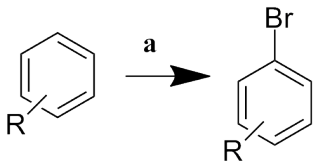
C-C bond formation is one of the most popular reactions in the 'medicinal chemist's toolbox', accounting for 11.5% of total transformations in a data set constituted of 7315 reactions and 3566 test compounds described in 139 publications as carried out by Roughley and Jordan.[471] Of these reactions, the most popular methodologies for achieving this transformation included the Suzuki-Miyaura cross-coupling reaction, Sonogashira reaction and as well as other palladium-mediated cross-coupling reactions.[472,473] Consequently, the preparation of privileged intermediates bearing functionalities reactive in these reactions, such as aryl bromides, has been of synthetic interest to the pharmaceutical chemistry community.

Aryl bromides can be prepared from an aniline using the Sandmeyer reaction (Scheme 4.3).[474] Diazotization of the aniline to form a diazonium salt, followed by subsequent attack by a nucleophile under copper catalytic conditions, yields an aryl bromide.



Scheme 4.3: Sandmeyer reaction for formation of an aryl bromide from an aniline via formation of a diazonium salt.(a) HBr, $NaNO_2$, 0 °C; (b) CuBr, 60-100 °C.

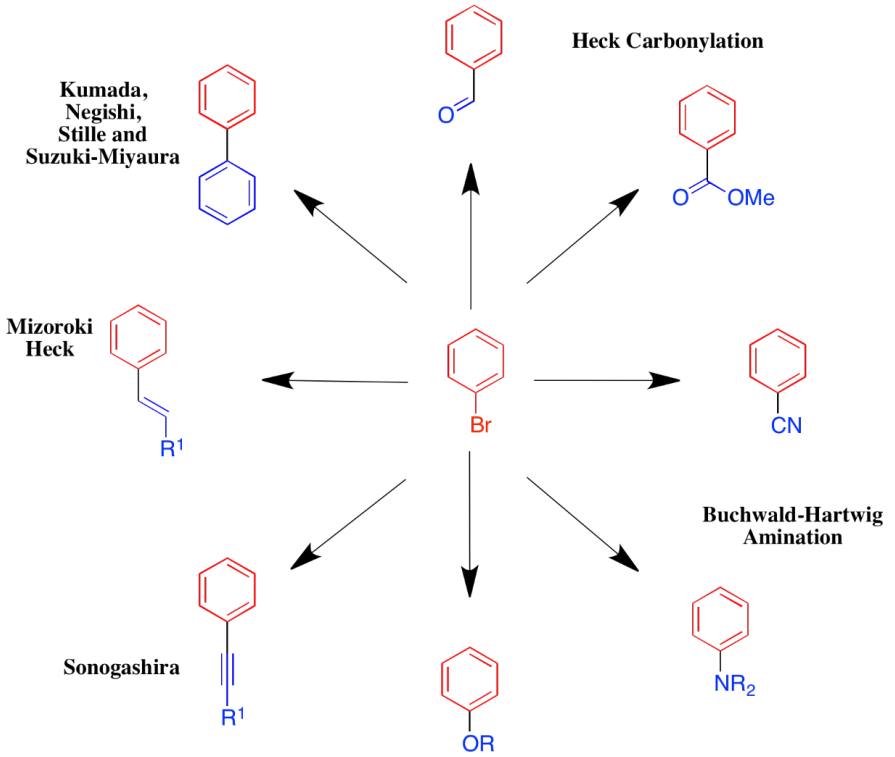
Alternatively, an electrophilic aromatic substitution reaction can be used to directly halogenate the aromatic ring (Scheme 4.4).[475] Typically, a Lewis acid catalyst is required to form a more reactive highly electrophilic complex with bromine. The regiochemical outcome can be influenced through the use of a directing group in order to achieve the desired substitution pattern. Furthermore, when the directing group is electron-donating, a catalyst is not necessary.



Scheme 4.4: Electrophilic aromatic bromination of an aromatic ring where R is electron donating. (a) $HBr_{(aq)}$, DMSO, EtOAc, 60 °C, 5 min – 8 h.

However, due to the high chemoselectivity of poly-halogenated aromatic systems in palladium-catalysed cross-coupling reactions, as well as their relative inertness to a vast array of nucleophilic and electrophilic reaction conditions (exceptions for strong nucleophiles), and the wide availability of a large range of halogenated aromatic starting materials, installation of the halogen typically occurs through careful selection of starting material.

Despite their relative inertness, aryl bromides are undoubtedly synthetically versatile, showing reactivity in Grignard reactions, lithium-halogen exchange reactions, benzyne formation as well as reactivity towards a wide variety of metal-mediated cross-coupling reactions yielding a diverse group of products (Scheme 4.5).[469]

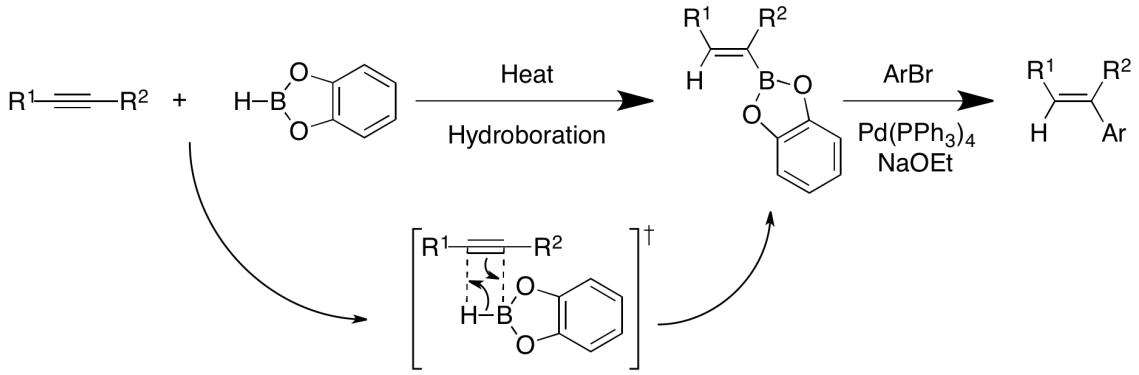


Scheme 4.5: Examples of different Palladium-catalyzed cross-coupling reactions of aryl bromides.

4.4.1 The Suzuki-Miyaura Reaction

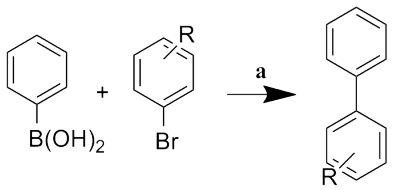
The Suzuki-Miyaura reaction (SM) was first published in 1979 by Akira Suzuki and has become of major importance due to the ubiquity of aromatic and heteroaromatic moieties in fine chemical intermediates, pharmaceuticals, agrochemicals and new materials.[476] The success of the reaction can be attributed to the wide availability of starting materials, broad tolerance of palladium catalysts towards various functional groups, and the reliable reproducibility of the results, even on an industrial scale. Accordingly, Suzuki shared the 2010 Nobel Prize for Chemistry with Heck and Negishi for the development of palladium-catalysed cross-couplings in organic synthesis.

The SM reaction was first exemplified by the hydroboration of an alkyne with a catecholborane followed by palladium(0)-catalysed coupling of the resulting vinyl boronate with an aromatic iodide or bromide (Scheme 4.6).[477] The hydroboration is generally regioselective for the less hindered position and addition of boron and hydrogen occurs *cis*-stereospecifically.



Scheme 4.6: First example of a Suzuki-Miyaura palladium-catalysed cross-coupling reaction.

The SM reaction describes the palladium-catalysed cross-coupling reaction between an aryl bromide and an organoboronic acid (Scheme 4.7). However, this can be extended to include the coupling of not just aromatic halides by also alkyl, alkenyl and alkynyl halides as well as the coupling of some pseudohalides such as triflates.[472] Likewise, alternative organoboron compounds such as, potassium trifluoroborates, organoboranes, and boronate esters such as MIDA boronates further expand the selection of available starting materials, where alternative boron reagents either offer superior bench stability or enhance reactivity over organoboronic acids.[478]

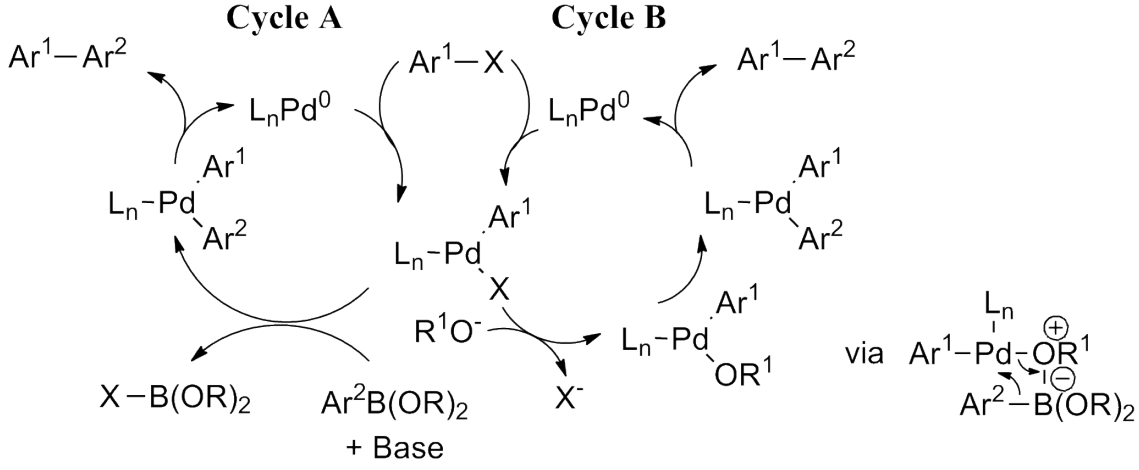


Scheme 4.7: Suzuki reaction of a boronic acid with an aryl bromide. (a) 2 eq $K_2CO_{3(aq)}$, 3 mol% $Pd(PPh_3)_4$, benzene, Δ.

The mechanism of the reaction is believed to operate by oxidative addition of the halide to the palladium(0) complex which generates a palladium (II) intermediate. This then undergoes transmetallation with the boronate coupling partner, from which the product is formed via reductive elimination, regenerating the palladium (0) catalyst (Scheme 4.8).[479]

Oxidative addition is often the rate-determining step in the catalytic cycle with a relative order of reactivity I > OTf > Br >> Cl. Organohalides in proximity to electron-withdrawing groups are more reactive than those bearing electron-donating groups.[472]

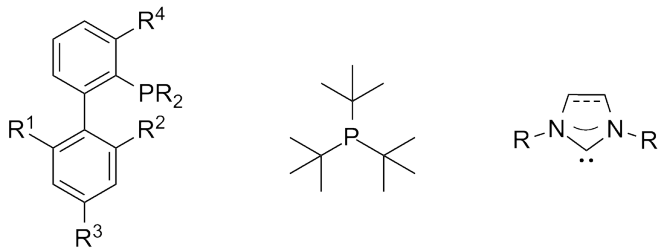
It is commonly observed that the rate of the transmetallation step is enhanced by the addition of base, typically sodium or potassium ethoxide or hydroxide. The reason behind this enhancement is commonly disputed. It is attributed to either facilitating the slow transmetallation step by formation of a more reactive boronate species that can interact with the Pd center and transmetallate in an intramolecular fashion (Scheme 4.8, cycle A),[480] or, the base may displace the halide in the coordination sphere of the palladium complex, facilitatating an intramolecular transmetallation (Scheme 4.8, cycle B).[481]



Scheme 4.8: General mechanism of the SM reaction.

Until around 1998, the SM reaction was conducted almost exclusively by employing a triarylphosphine ligand. However, more recently the development of alternative ligands

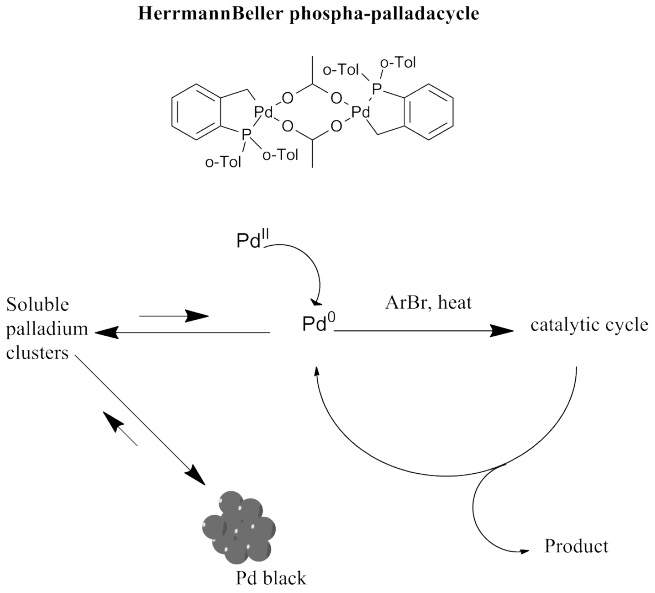
has improved the selectivity and efficiency of these reactions (Figure 4.8). For example, bulky dialkylbiaryl phosphines,[482,483] trialkylphosphine,[484] and *N*-heterocyclic carbenes.[485]



Dialkylbiaryl phosphine **Trialkylphosphine** **NHC's**

Figure 4.8: Chemical structures of modern ligands for the SM reaction.

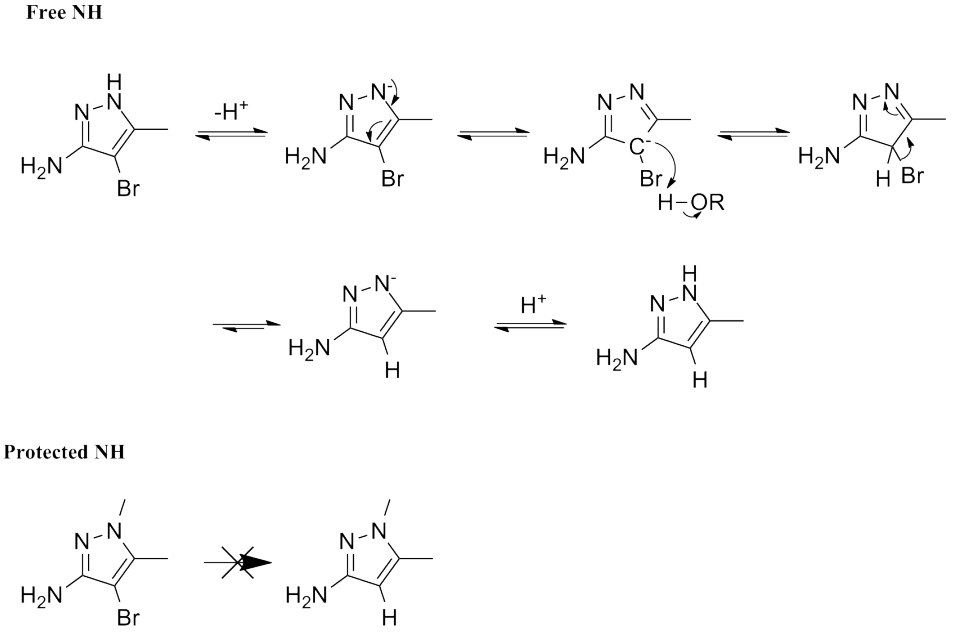
In most cases the exact nature of the catalyst remains ambiguous, varying from a L_2Pd^0 species to a highly reactive mono-ligated LPd species, where the L:Pd ratio can play a large role in catalytic performance.[486] However, in some cases, such as for palladacycle catalysts, the active catalyst is thought to be ligand-free palladium(0), where the palladacycle serves as a thermally stable source of ligand-free palladium(0), allowing harsher conditions and hence the use of cheaper, less reactive substrates.[487] The high turnover numbers observed were rationalized in terms of a pre-equilibrium forcing the reaction away from the formation of palladium black (Scheme 4.9).



Scheme 4.9: Pre-equilibrium describing the liberation of ligand-free palladium from a palladacycle catalyst.

Despite the remarkable synthetic utility of the SM reaction, there remains some coupling partners that pose major problems, such as, unactivated aryl chlorides, heteroaromatics and more hindered systems.[488–490] For example, heterocyclic substrates hinder the SM

reaction by co-ordination of the heterocycle substrate or product to the metal centre resulting in deactivation of the catalyst.[491,492] Furthermore, instability and/or poor reactivity of some heterocyclic boronic acids leading to side reactions such as proto-deborylation or oxidative homocoupling or dehalogenation.[493] Fortunately, the development of new ligands and catalysts has further expanded the scope of coupling partners that can be employed in the SM reaction to include some heterocyclic substrates.[494–498] However, electron-rich heterocyclic substituents, such as those bearing free NH groups, for example pyrazoles or imidazoles, are still not tolerated due to the low reactivity of the oxidative addition complex, inhibition of the Pd(II) starting material or instability of the starting material leading to decomposition, for example, via dehalogenation (Scheme 4.10).[499,500]



Scheme 4.10: Proposed dehalogenation of pyrazole.

However, catalyst deactivation and dehalogenation is not observed for heterocycles that are functionalized at NH, meaning that by the use of protecting group chemistry, this limitation can be overcome.[501]

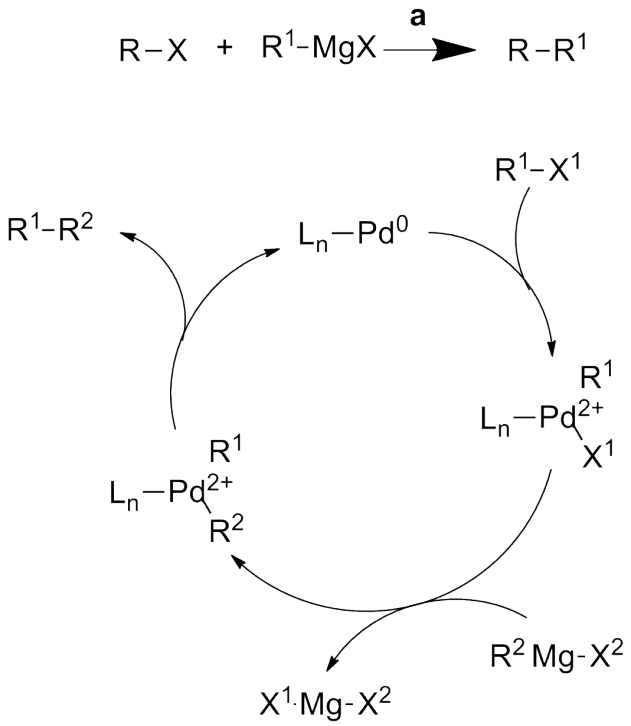
4.4.2 Other Palladium-Catalyzed Cross-Coupling Reactions

Additionally to the SM reaction, aryl bromides are reactive to a plethora of other palladium-catalyzed cross-coupling reactions such as, but by no means limited to, the Kumada, Negishi, and Stille, which, similarly to the SM reaction, can yield biaryls, as

well as the Heck, Sonogashira and Buchwald-Hartwig cross-coupling reactions that yield aromatic alkenyl, alkynyl and amines respectively.[469]

4.4.2.1 The Kumada Reaction

The Kumada coupling reaction was first reported by Kumada and Corriu independently in 1972 and is the palladium or nickel catalysed cross-coupling reaction of an alkyl, vinyl or aryl halide and a Grignard reagent to form a new carbon-carbon bond (Scheme 4.11).[502–504] The Kumada coupling was the first example of metal-catalyzed carbon-carbon bond formation that inspired the development of many related reaction such as the SM and the Negishi in the following years.[505]

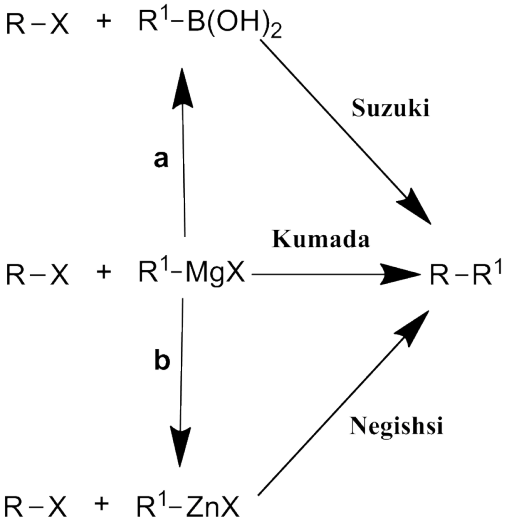


Scheme 4.11: General Kumada coupling reaction. (A) General reaction scheme for a Kumada coupling (a) $Pd(PPh_3)_4$ or $Ni(dppb)Cl_2$, solvent, Δ. (B) General catalytic cycle for Kumada coupling of an organohalide and a Grignard reagent.

4.4.2.2 The Negishi Reaction

The Negishi coupling, discovered in 1977, is very closely related to the Kumada coupling reaction, as shown in Scheme 4.12.[506] The coupling occurs between an organozinc compound with an organohalide. The reaction is not restricted to the formation of unsymmetrical biaryls and can be extended to the coupling of alkenyl, aryl, allyl, benzyl, and propargyl groups and hence is known for its broad functional group tolerance. The

following reactivity for palladium-catalyzed cross-coupling reactions according to the nucleophilic substrate is Zn > Mg > Li where challenging products can be most often obtained by the Negishi reaction.[507]



Scheme 4.12: Relationship between the Kumada reaction and the Suzuki and Negishi reactions. (a) $B(OMe)_3$ then $H_2SO_4(aq)$; (b) $ZnBr_2$.

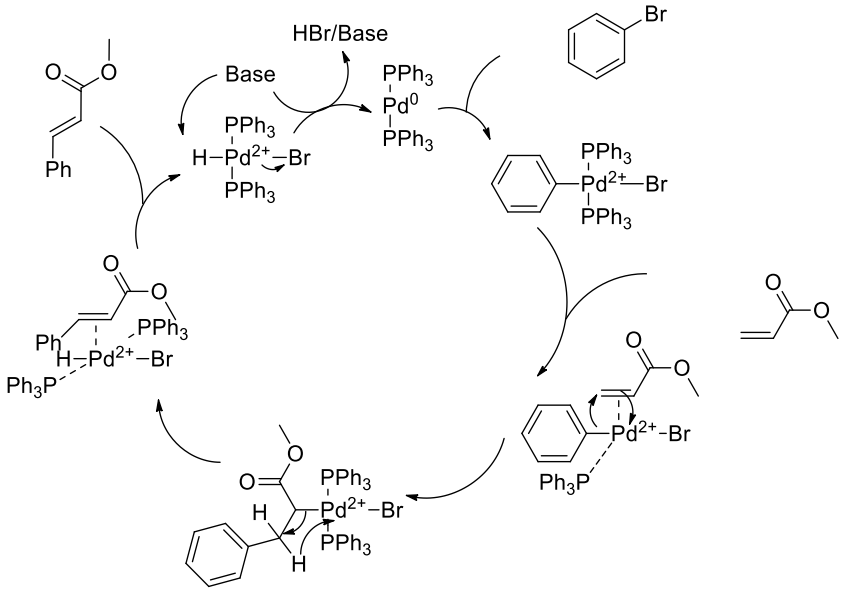
4.4.2.3 The Stille Reaction

The Stille coupling is yet another C-C bond forming reaction between stannanes and halides or pseudohalides that has, similarly to the Negishi coupling, very few limitations of the nature of R groups that can be coupled.[508,509] It follows the same general mechanism as that depicted in Scheme 4.8, that follows an oxidative addition, transmetallation, *trans/cis* isomerization follow by reductive elimination. The major drawback of the Stille reaction is the toxicity of the tin compounds used as well as their poor solubility in water.[510] Large advances in SM chemistry, associated with development of ligands that show higher efficiency and selectivity as well as the introduction of microwave-assisted organic synthesis has led to a decline in the use of the Stille reaction as the organoboron coupling partners are not toxic and are reasonably soluble in water.

4.4.2.4 The Heck Reaction

The Mizoroki-Heck (commonly known as Heck) reaction was developed in 1971 and is the palladium-catalyzed cross-coupling reaction that occurs between an aryl halide and a vinyl halide, pseudohalide or activated halide (such as methyl acrylate) in the presence of

a base.[511,512] The Heck reaction proceeds under mild conditions with excellent *trans*-selectivity and is amenable to scale-up procedures.



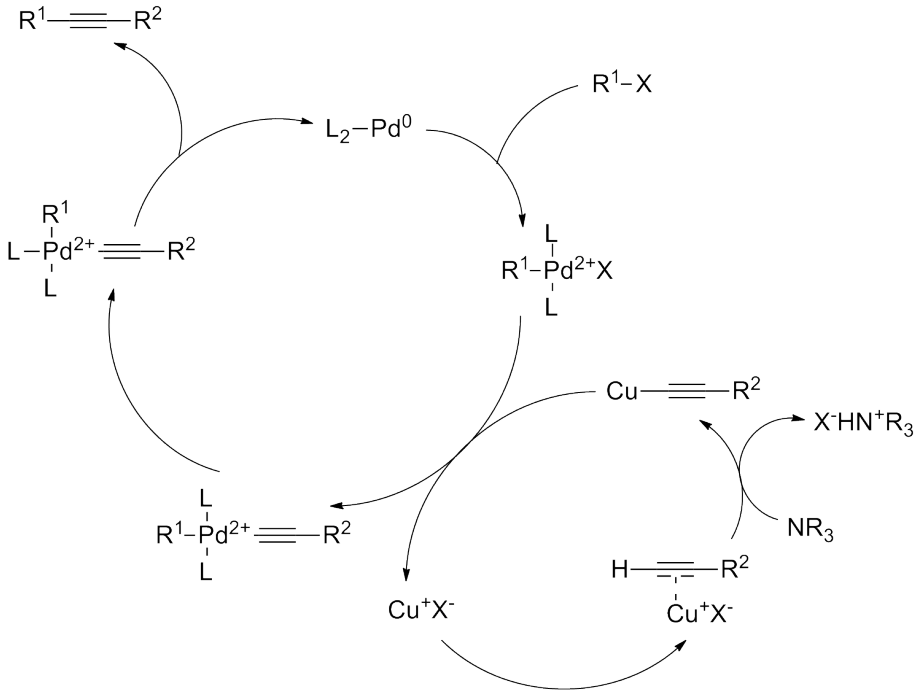
Scheme 4.13: Catalytic cycle of the Heck reaction.

The mechanism of the Heck reaction starts by oxidative addition of the aryl halide followed by insertion of the vinylic coupling partner to form a Pd-σ intermediate that, upon β-hydride elimination forms a π-complex that eliminates product. Base then facilitates reformation of catalytically active Pd^0 species (Scheme 4.13).

4.4.2.5 The Sonogashira Reaction

The Sonogashira reaction is the palladium-catalyzed cross-coupling reaction of terminal alkynes with aryl or vinyl halides in the presence of a copper co-catalyst and an amine base to yield acetylenic compounds.[513] The mechanism is not fully understood, with confusion arising from the combined action of two metal catalyst, but is widely believed to take place through two independent catalytic cycles, corresponding to palladium and copper catalysed transformations.[514] The palladium-mediated cycle is believed to be analogous to the 'classical' palladium-mediated cross-coupling reactions, constituting of oxidative addition of the organohalide, transmetallation of the acetylenic moiety followed by isomerization to allow reductive elimination to yield the product and catalytically active Pd^0 (Scheme 4.14).[515] By contrast, the copper cycle is poorly known. The base is believed to assist in copper acetylide formation via a π-alkyne copper complex (similar to Heck). The resultant copper acetylide then transmetallates into the palldium cycle (Scheme 4.14). However, this has been complicated by the observation of the formation

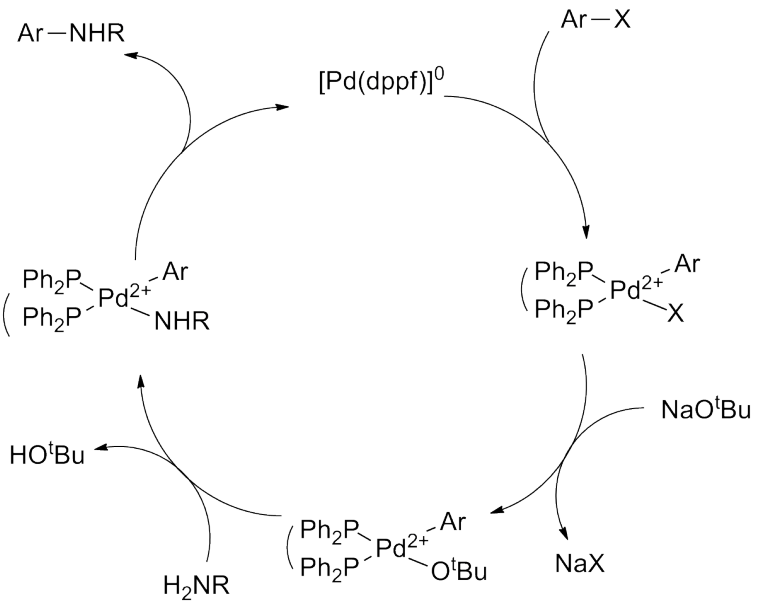
of CuI-polyphosphane adducts, showing that copper-ligand interactions could also be likely.[516,517] Furthermore, the role of the amine has be theorized to potentially be multi-faceted, with involvement not only in deprotonation of the alkyne complex but also interfering in oxidative addition, resulting in the formation of a more reactive [Pd^0L(amine)] complex.[518]



Scheme 4.14: General catalytic cycles for the Sonogashira reaction.

4.4.2.6 The Buchwald-Hartwig Reaction

The Buchwald-Hartwig reaction uses palladium-mediated, cross-coupling reactions to synthesise anilines from aryl halides and amines (Scheme 4.15).[519,520] However, this may be extended to include the synthesis of ethers by the replacing the amine with a phenol providing an alternative to copper-catlayzed Ullmann and Goldberg couplings.[521] Functional group tolerance is higher than for copper-catalyzed methods, allowing the coupling of heteroaromatic halides in good to excellent yields.



Scheme 4.15: Catalytic Cycle of the Buchwald-Hartwig amination reaction.

4.4.3 Summary

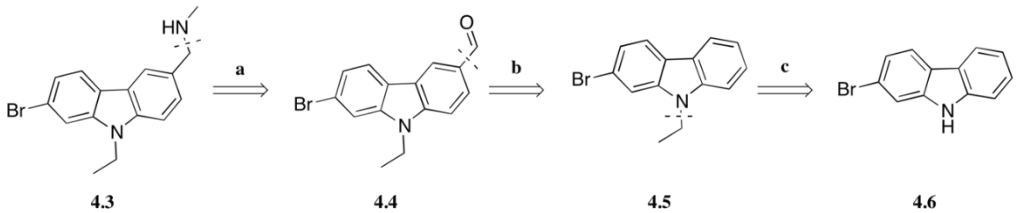
In general, aryl bromides are reactive to a wide range of palladium-catalysed cross-coupling reactions, leading to the formation of new carbon-carbon, carbon-nitrogen or carbon-oxygen bonds depending on the conditions employed.

Hence, by synthesis of the privileged intermediate **4.3**, a diversity-oriented synthetic strategy can be utilized, to probe the steric and electronic limits of subsite II, seeking optimize PK083 analogues toward p53-Y220C.

4.5 Results and Discussion

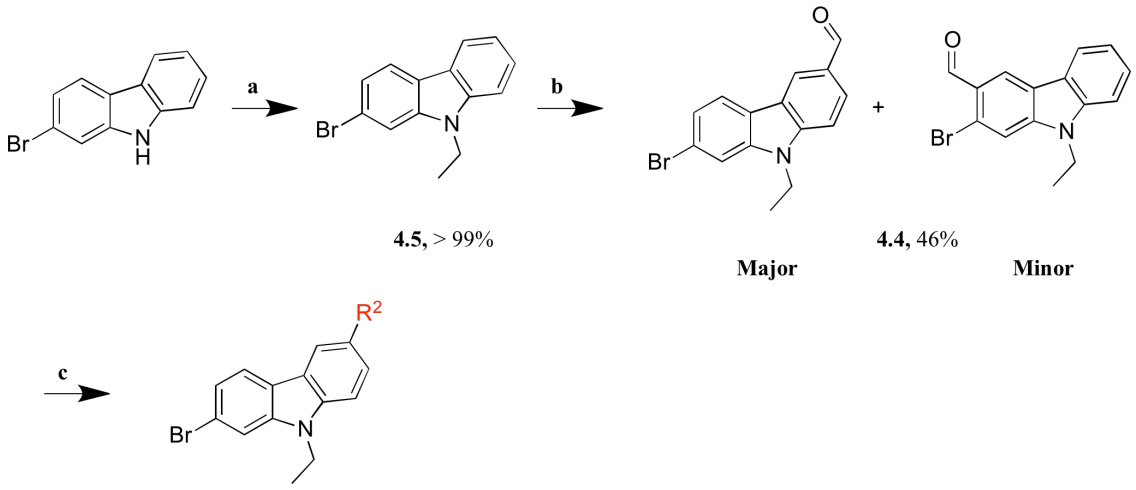
4.5.1 Synthesis of Intermediate **4.3**

The synthetic strategy towards intermediate **4.3** is summarized by the retrosynthesis described in Scheme 4.16.



Scheme 4.16: Reterosynthesis of **4.3**. (a) $H_2NMe{\cdot}HCl$, Et_3N, $Ti(O^iPr)_4$, DCM/EtOH, rt 18 h then $NaBH_4$, rt, 8 h; (b) $POCl_3$, DMF, Δ; (c) NaH, THF, rt then CH_3CH_2I.

The ethylation of 2-bromocarbazole yielded 2-bromo-9-ethyl-9*H*-carbazole in an excellent yield. Vilsmeier-Haack formylation of 2-bromo-9-ethyl-9*H*-carbazole yielded two regio-isomeric products, 7-bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde and 2-bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde, where the desired product was the major isomer formed.[470] Reductive amination with both methylamine and dimethylamine yielded 2-bromo PK083 and PK8017 analogues respectively (Scheme 4.17).[367]

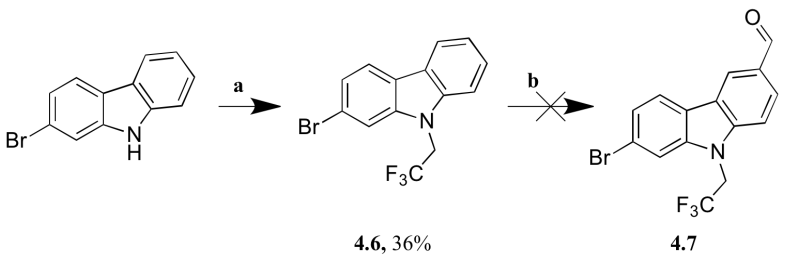


Scheme 4.17: Synthesis of **4.3a** and **4.3b**. (a) NaH, DMF, rt, 30 min then CH_3CH_2I; (b) $POCl_3$, DMF, 70 °C 18 h; (c) amine·HCl, Et_3N, $Ti(O^iPr)_4$, DCM/EtOH, rt 18 h then $NaBH_4$, rt, 8 h.

4.5.2 Synthesis of a trifluoroethylated intermediate

The two privileged intermediates, **4.3a** and **4.3b**, constitute the brominated analogues of PK083 and PK8017. However, application of this three-step synthesis to the synthesis of a 2-bromominated analogue of the most potent PK083 analogue described so far, PK9255, that bears an *N*-trifluoroethylated functionality, was unsuccessful (Scheme 4.18).

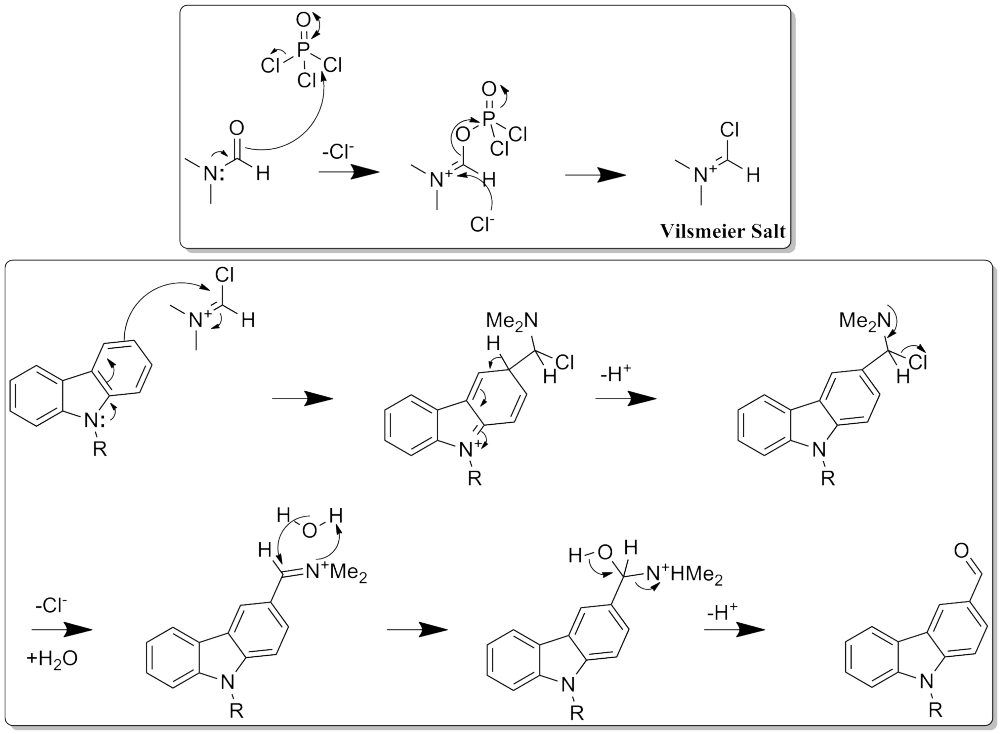
Trifluoroethylation of 2-bromocarbazole according to the procedure established in chapter 3 afforded 2-bromo-9-(2,2,2-trifluoroethyl)-9*H*-carbazole in a moderate yield. Vilsmeier-Haack formylation of 2-bromo-9-(2,2,2-trifluoroethyl)-9*H*-carbazole was unsuccessful under a variety of conditions (higher equivalents of Vilsmeier salt and microwave conditions) (Scheme 4.18).



4.6, 36% **4.7**

Scheme 4.18: Attempted synthesis of **4.7**. (a) $TsOCH_2CF_3$, Cs_2CO3_3, DMF, μwave, 150 °C, 30 min; (b) $POCl_3$, DMF, Δ.

The Vilsmeier-Haack formylation reaction is the reaction between an electron-rich aromatic group and the Vilsmeier salt, which, upon work up results in the formation of an aldehyde according the mechanism described in Scheme 4.19.

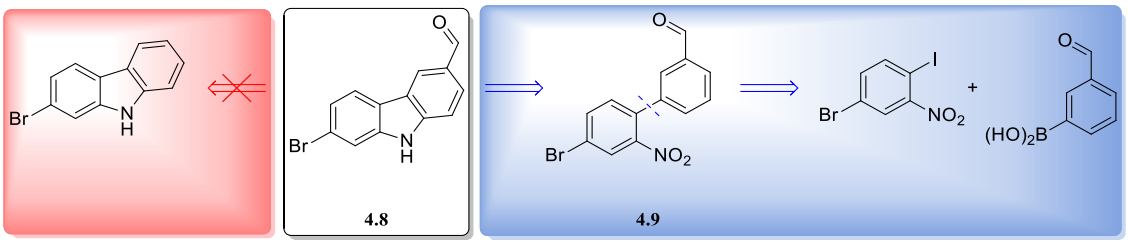


Scheme 4.19: Mechanism of Vilsmeier-Haack formylation.

The powerful electron-withdrawing effect of the trifluoroethyl group of **4.6** versus the ethyl group of **4.5** may inductively prevent the lone pair of the nitrogen of carbazole from delocalizing in the ring to attack the Vilsmeier salt according to the mechanism described in Scheme 4.19.

Synthetically this may be overcome by installation of the aldehyde prior to N-functionalization. This cannot be accomplished simply by Vilsmeier-Haack on 2-bromocarbazole, as carbazole's bearing a free NH group have been shown to exhibit a

preference for N-formylation, conversely to the reactivity exhibited by indole (Section 2.5.1, Scheme 2.4). Therefore, synthesis of the required intermediate **4.7** may be accomplished through reductive cyclization of the biaryl described in Scheme 4.20, which in turn can be accessed through the Suzuki-Miyaura reaction of the two commercially available starting materials, 4-bromo-1-iodo-2-nitrobenzene and 3-formylphenyl boronic acid according to a literature procedure exemplified in Table 4.1.

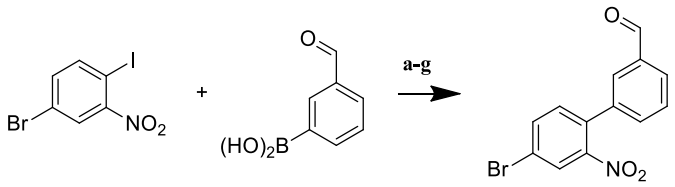


Scheme 4.20: Possible routes to intermediate **4.8**. Red, Vilsmeier-Haack of 2-bromocarbazole (not possible). Blue, reductive cyclization of the biaryl **4.9** formed by a Suzuki reaction between the building blocks 4-bromo-1-iodo-2-nitrobenzene and 3-formyl boronic acid.

Following the literature conditions established by Chaitanya *et al*., the nitrobiphenyl was obtained in a poor yield (conditions A, Table 4.1).[470] Reaction optimization, by variation of the catalyst and the conditions, resulted in conditions E, which gave the desired biaryl **4.9** in an 84% yield. Note, the more reactive C-I bond couples, but not the C-Br.

Reductive cyclization of intermediate **4.9** under classical Cadogan conditions i.e. solventless conditions using triethylphosphite, did not yield the desired product under reflux or microwave irradiation.[357] However, using triphenylphosphine in DMA under reflux conditions yielded the product as a mixture of regioisomers (Scheme 4.22).[522]

Failure of the 'classical' reagent, triethyl phosphite, to be effective for the cyclization was attributed to the fact that it can react with the aldehyde in a manner similar to dimethyl acetal aldehyde protection (Scheme 4.23).

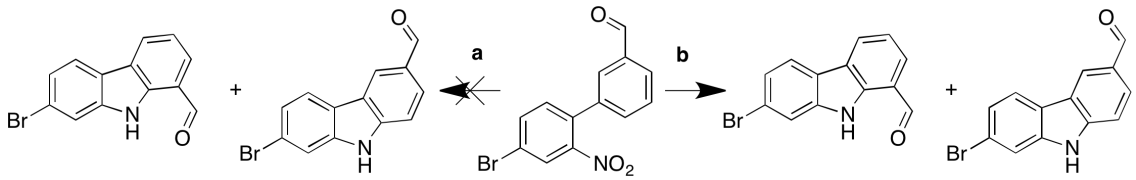


4.9

Scheme 4.21: Synthesis of 4.9.

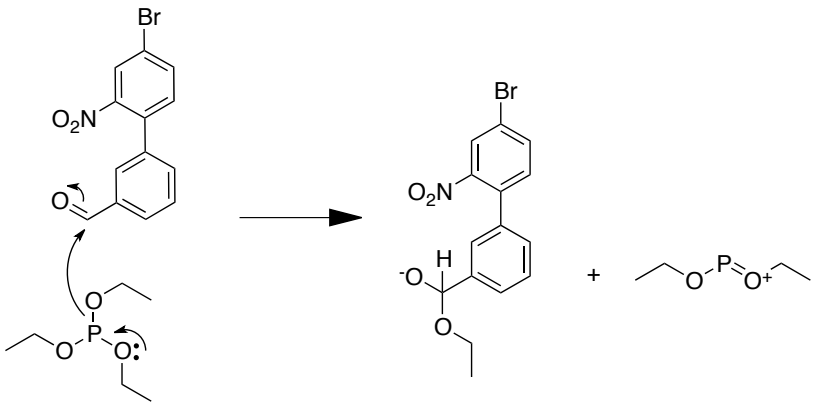
Entry	Catalyst	Base	Solvent	Conditions (°C)	Yield 4.9 (%)
A[a]	$Pd(OAc)_2$, PPh_3	K_2CO_3	Toluene, H_2O	70	6
B	$Pd(PPh_3)_4$	K_2CO_3	Toluene, H_2O	70	3
C	$Pd(PPh_3)_4$	K_2CO_3	Toluene, H_2O	150 (μwave irradiation)	12
D	$Pd(dppf)Cl_2.CH_2Cl_2$	K_2CO_3	Toluene, H_2O	70	46
E	$Pd(dppf)Cl_2.CH_2Cl_2$	K_2CO_3	Toluene, H_2O	90	84
F	Pd-118	K_2CO_3	Toluene, H_2O	70	0
G	Pd-118	K_2CO_3	Toluene, H_2O	25	18

Table 4.1: Reaction optimization of the Suzuki-Miyaura reaction described in Scheme 0.15. aConditions according to Chaitanya et al.[470]



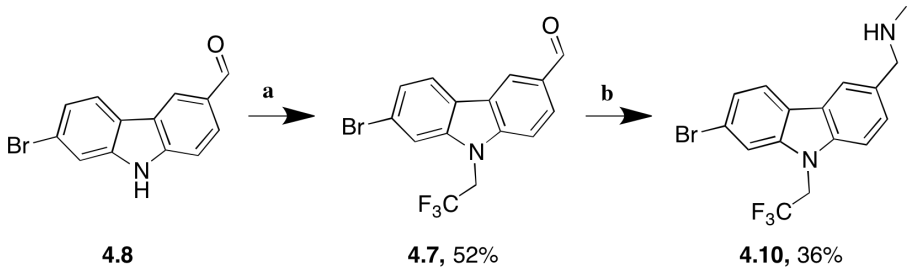
4.8 **4.9** **4.8,** 22%

Scheme 4.22: Reductive cyclization of **4.9** (a) $P(OEt)_3$, Δ; (b) PPh_3, DMA, 165 °C, 18 h.



Scheme 4.23: Proposed mechanism for reaction of triethyl phosphite with 4.9.

Trifluoroethylation of **4.8** yielded the desired product, which, when subjected to reductive amination, yielded the intermediate **4.10**, that constitutes the 2-bromo analogue of PK9255 (Scheme 4.24).



4.8 **4.7,** 52% **4.10,** 36%

Scheme 4.24: Synthesis of **4.10**. (a) $TsOCH_2CF_3$, Cs_2CO_3, DMF, μwave, 150 °C, 30 min; (b) $H_2NMe.HCl$, $Ti(O^iPr)_4$, NEt_3, DCM, EtOH rt 18 h then $NaBH_4$, rt, 8 h.

4.5.3 Synthesis of a small library of intermediates via Cadogan Cyclization

Advantages of using the Cadogan synthesis, using triphenylphosphine as a reagent, include the high functional group tolerance of the substituents around the nitrobiphenyl.[522] Accordingly, this expands the choice of reactive groups that form intermediates that can be used to target subsite II at position R^3 (Figure 4.9). Hence, the choice of intermediates may be greater expanded in order to further enhance the diversity possible by reaction at R^3.

The reactive groups at R^3 were selected for their differing reactivity versus an aryl bromide. Table 4.2 summarizes the expansion of intermediates and their associated reactivity that allows for the diverse reactivity at R^3.

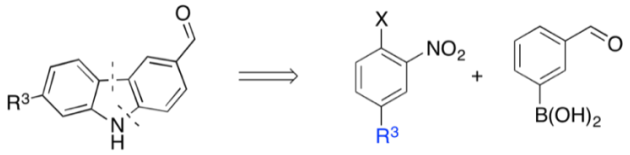
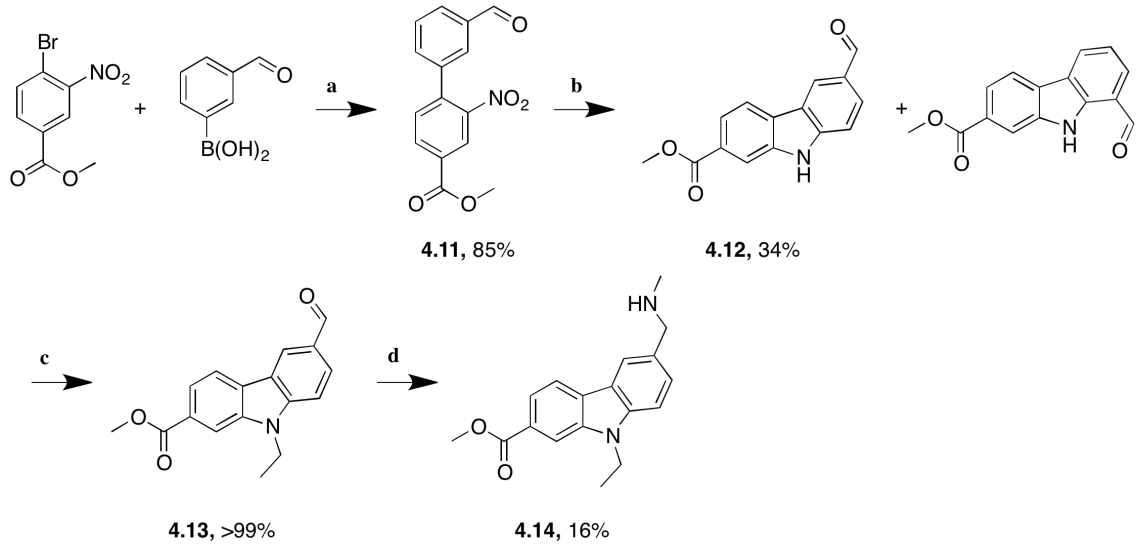


Figure 4.9: Retrosynthetic disconnection of a generic intermediate to commercially available building blocks.

R^3	X	Reaction
OH (**4.19**)	Br	Ullmann ether synthesis, Williamson ether synthesis
Br (**4.8**)	I	Pd-catalyzed cross-coupling reactions, benzyne formation, Grignard and lithiation reactions
MeO_2C (**4.12**)	Br	Amide synthesis, nucleophilic substitution reactions

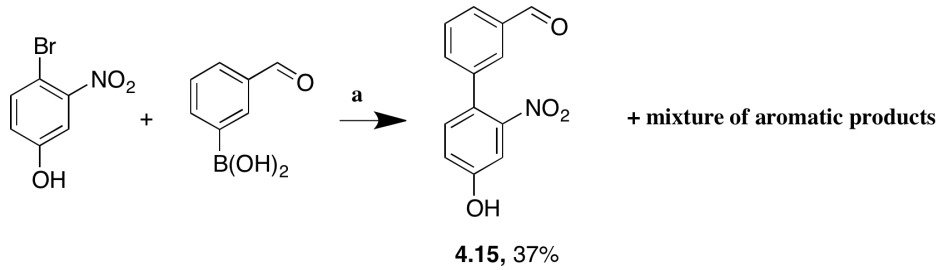
Table 4.2: Comparison of the nature of the nitro building block at X and R^3 and the associated reactivity of the resulting carbazole intermediate.

Using the optimized conditions developed in the synthesis of **4.10**, the synthesis of compounds **4.12** and **4.19** was attempted.



Scheme 4.25: Synthesis of compound **4.14** via formation of intermediate **4.12**. (a) $Pd(dppf)Cl_2.CH_2Cl_2$, K_2CO_3, toluene, H_2O, 90 °C, 1 h; (b) PPh_3, DMA, reflux; (c) NaH, DMF, rt, 30 min then CH_3CH_2I, 18 h; (d) $H_2NMe.HCl$, $Ti(O^iPr)_4$, NEt_3, DCM, EtOH rt 18 h then $NaBH_4$, rt, 8 h.

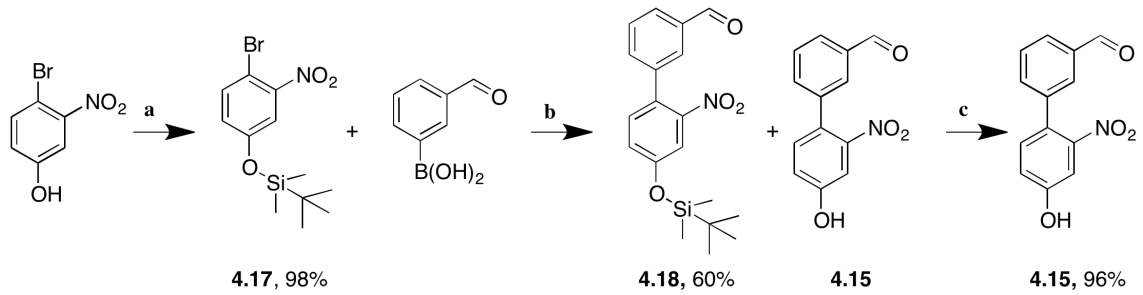
The Suzuki-Miyaura cross coupling reaction between 4-bromo-3-nitrobenzoate and 3-formylphenyl boronic acid yielded the nitrobiphenyl **4.11** in an 85% yield, Reductive cyclization yielded the product in a 34% yield as a mixture of regioisomers. Ethylation followed by reductive amination yielded the compound **4.14** (Scheme 4.25).



Scheme 4.26: SM reaction of 4-bromo-3-nitrophenol. (a) $Pd(dppf)Cl_2.CH_2Cl_2$, K_2CO_3, toluene, H_2O, 90 °C.

Conversely, initial attempts at the synthesis of the phenol intermediate **4.15** were unsuccessful. The Suzuki-Miyaura cross-coupling between 4-bromo-3-nitrophenol and 3-formylphenyl boronic acid was low yielding (37% product formation) and led to an inseparable complex mixture of products (Scheme 4.26). Phenols are reactive in metal-catalyzed cross-coupling reactions, for example, the Chan-Lam (phenol and boronic acid) and Ullmann diaryl ether synthesis (phenol and bromide). Protection of the phenol to a *tert*-butyldimethylsilyl ether **4.17** limited the side reaction and yielded a silyl protected biaryl **4.08** as well as the unprotected phenol biaryl **4.15** (Scheme 4.27). Cleavage of the

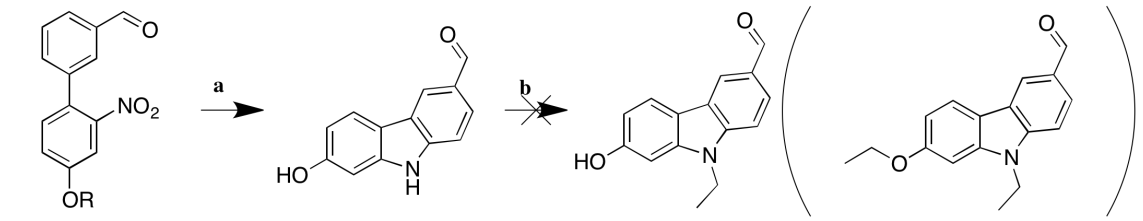
silyl ether is postulated to occur thermolytically after the completion of the Suzuki reaction upon sustained heating (overnight).



4.17, 98% **4.18,** 60% **4.15** **4.15,** 96%

Scheme 4.27: Silyl protection of 4-bromo-3-nitrophenol. (a) TBDMSCl, Et_3N, DCM, 0 – 20 °C (b) $Pd(dppf)Cl_2.CH_2Cl_2$, K_2CO_3, toluene, H_2O, 90 °C; (c) TBAF, THF, 0 °C – rt, 20 min.

Reductive cyclization of both **4.18** and **4.15** afforded the carbazole **4.19**, where desilylation was proposed to occur thermolytically under the reductive cyclization conditions. Chemoselective N-ethylation was attempted, using conditions reported for ethylation of 2-hydroxycarbzole, however, these conditions failed to yield the desired N-ethylated free phenolic product, favouring the diethylated carbazole **4.20** (Scheme 4.28).

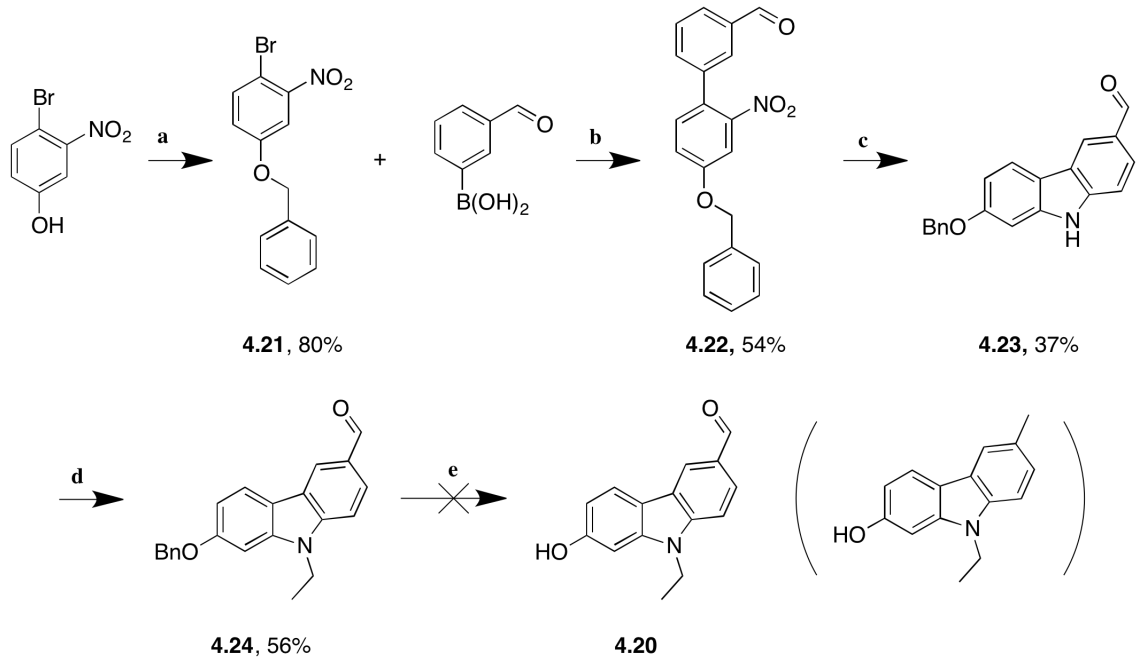


R = H (**4.15**)
R = TBDMS (**4.18**)

4.19, 21% **4.20,** 17%

Scheme 4.28: Reductive cyclization of **4.15** and **4.18** yields **4.19**. (a) PPh_3, DMA, reflux; (b) NaH, DMF, THF, rt, 10 min, then CH_3CH_2I, THF, dropwise.

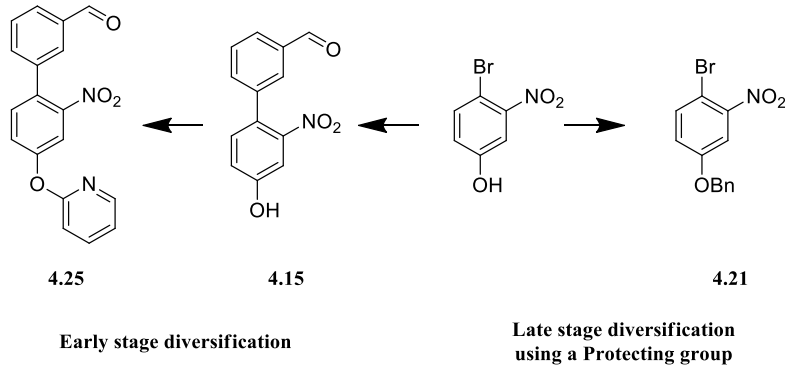
Whilst protection of the phenol to a silyl ether provides an adequate solution for increasing the yield of **4.15**, it does not offer a solution for the lack of chemoselectivity in the synthesis of **4.20** due to thermolytic cleavage of the silyl ether under the harsh conditions required for reduction cyclization. Thus, a more thermally stable, robust benzyl protecting group was tested for its amenability to these optimized conditions (Scheme 4.29).



Scheme 4.29: Synthesis of a benzyl ether. (a) BnBr, NaH, DMF, rt; (b) $Pd(dppf)Cl_2.CH_2Cl_2$, K_2CO_3, toluene, H_2O, 90 °C; (c) PPh_3, DMA, reflux; (d) NaH, DMF, rt, 30 min then CH_3CH_2I; (e) Pd/C, H_2, THF, rt.

Whilst this strategy was both effective at minimizing side products in the Suzuki reaction, and circumventing chemoselectivity problems, the harsh deprotection conditions led to reduction of the aldehyde to a tolyl group.

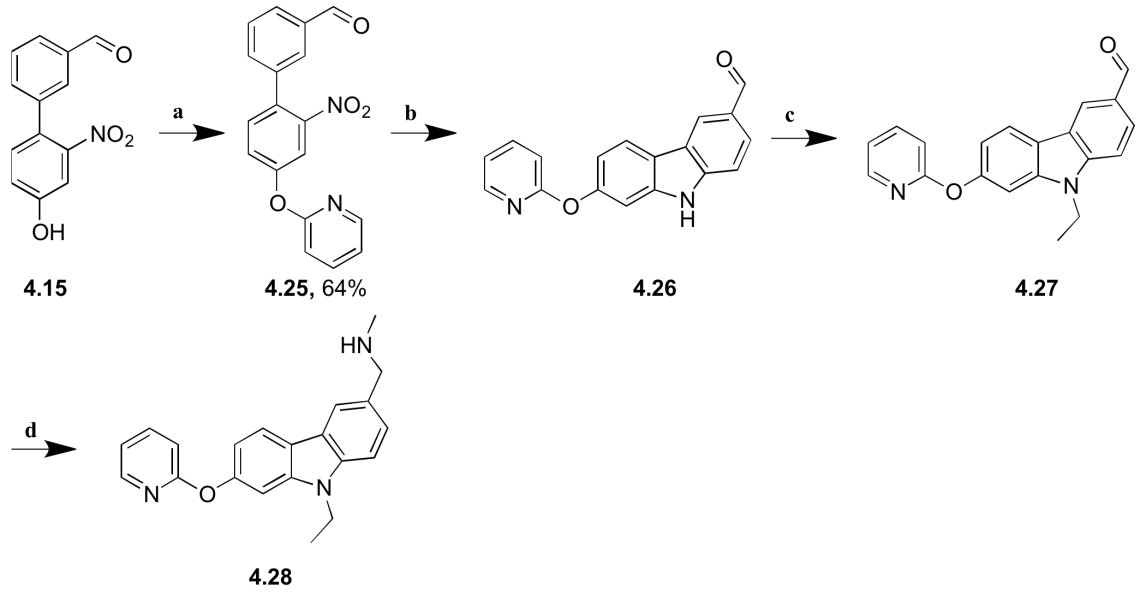
One method of circumventing this is to use early stage functionalization, i.e. installation of the subsite II group, through either a Williamson ether synthesis of Ullmann diaryl ether synthesis of the biaryl phenol intermediate **4.15**. This technique overcomes the lack of chemoselectivity in the attempted ethylation of **4.23** by installing the desired subsite II targeting group earlier in the synthesis before reductive cyclization (Scheme 4.30).



Scheme 4.30: Comparison of early-stage and late-stage funcitonalization techniques.

Ullmann coupling of biaryl **4.15** proceeded in a good yield to yield the early-stage functionalized pyridine ether **4.25**, which, upon reductive cyclization yielded the

carbazole **4.26**. N-Ethylation, followed by reductive amination, yielded the final product **4.28** (Scheme 4.31).



Scheme 4.31: Synthesis of ether product **4.28** by early-stage functionalization of intermediate **4.15**. (a) 2-Bromopyridine, 2-picolinic acid, Cu(I)I, K_3PO_4, DMSO, 100 °C, 3 h; (b) PPh_3, DMA, reflux; (c) NaH, DMF, rt, 30 min then CH_3CH_2I; (d) $H_2NMe.HCl$, $Ti(O^iPr)_4$, NEt_3, DCM, EtOH rt 18 h then $NaBH_4$, rt, 8 h.

From the above reactions, a small library of late-stage and one early stage intermediate was synthesised that allow diversification at position R^3 to target subsite II (Figure 4.10).

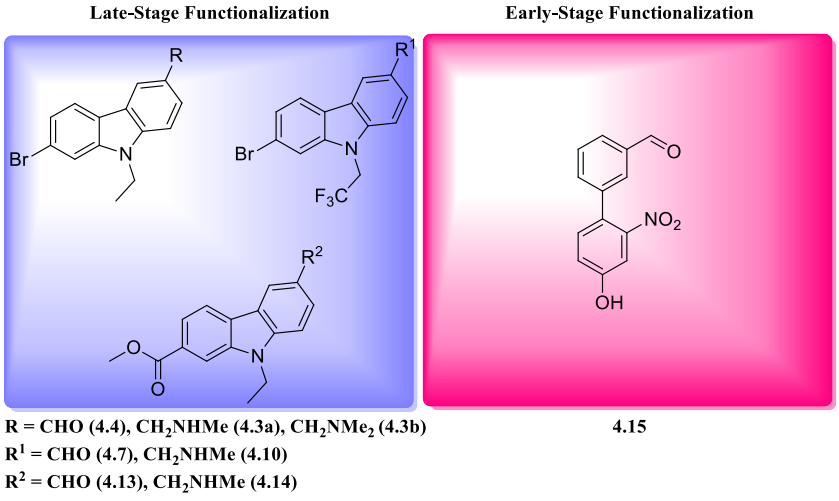
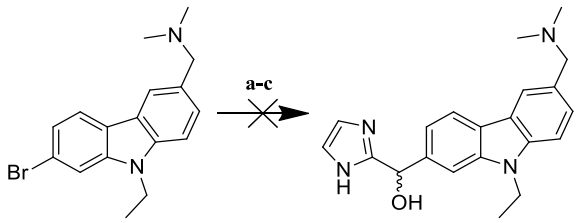


Figure 4.10: Summary of synthesised intermediates for either early or late stage functionalization.

4.5.4. Initial Diversification

Using the intermediates depicted in Figure 4.10, reaction conditions towards the generation of a structurally diverse library were sought. Late-stage intermediates were prioritized due to their amenability towards rapid library generation over early-stage diversification that is hindered by and bottlenecks such as repetitive chemistry, work-ups and purifications.

The reactivity of aryl bromides is summarized in Section 4.4, and includes a wide range of metal-catalyzed cross-coupling reactions as well as Grignard or lithiation chemistry. Initial attempts to use intermediate **4.3b** in a lithiation and Grignard reaction, under a range of conditions, to afford product **4.29** were unsuccessful (Scheme 4.32).



4.3b **4.29**

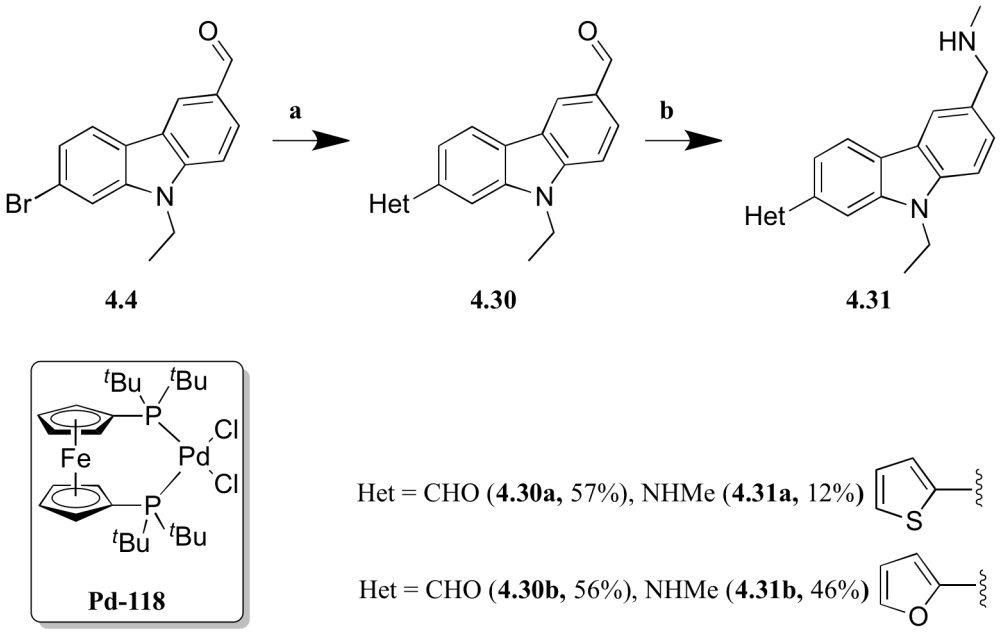
Scheme 4.32: Initial attempt at late-stage diversification of **4.3b** through Grignard and lithiation chemistry. (a) Mg, I_2, THF then 1-trityl-1*H*-imidazole-2-carbaldehyde, THF; (b) ^{i}PrMgCl.LiCl, THF, -40 °C then 1-trityl-1*H*-imidazole-2-carbaldehyde, THF, rt; (c) ^{n}BuLi, -78 °C, THF, 30 min then 1-trityl-1*H*-imidazole-2-carbaldehyde, THF, -78 °C – rt.

Both the 'classical' Grignard and turbo-Grignard (^{i}PrMgCl.LiCl) conditions failed to initiate, resulting in full recovery of starting material.[523] Lithiation of **4.3b** with ^{n}BuLi at -78 °C in THF followed by nucleophilic attack of the organo-lithiate with the aldehyde of 1-trityl-1*H*-imidazole-2-carbaldehyde led to an inseparable mixture (Scheme 4.32). The complexity of the latter was presumably due to competition with chelation-controlled ortho-lithiation assisted by the benzylic amine.

The unsuccessful attempts at utilizing intermediate **4.3b** in Grignard and lithiation chemistry led to the exploration of **4.4**'s reactivity toward metal-mediated cross-coupling reactions. Using intermediate **4.4**, which possesses an aldehyde rather than the benzylic amine of **4.3**, simplifies the purification of the amine product.

The SM cross-coupling reaction between intermediate **4.4** and heterocyclic-2-boronic acid's, according to the procedure established by Moseley *et al.*, yielded intermediates

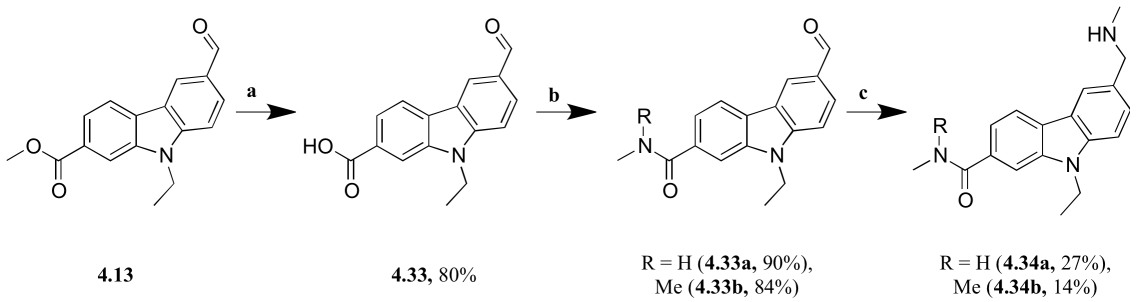
4.30a and **4.30b** in a 57 and 56% yield respectively. Reductive amination yielded the final products **4.31a** and **4.31b** (12%, 46% yields respectively) (Scheme 4.33).



Scheme 4.33: Diversification of **4.4** using a SM reaction with heterocyclic boronic acids. (a) thiophene-2-boronic acid or furan-2-boronic acid, Pd-118, K_2CO_3, MeCN, 90 °C, 2 h; (b) H_2NMe, $NaBH(OAc)_3$, THF, rt.

The SM reaction was also conducted under aqueous conditions and gave similar results. Pd-118, a highly active ferrocenyl phosphine palladium precatalyst, t that is commonly used to couple unactivated aryl chlorides, is sufficient to couple the electron-rich heteroaromatic boronic acids. The advantage of using Pd-118 over Pd(NHC) complexes is its superior stability to oxygen and water, whilst still being very active in SM reactions.[524]

Similarly to that conducted for aryl bromides, reaction conditions were sought with which to synthesise a diverse library based on the methyl ester intermediate **4.13**. Methyl esters are reactive to a plethora of reactions, such as amide couplings (via carboxylic acid) and nucleophilic substitution reactions. Synthesis of the related carboxylic acid was accomplished through base hydrolysis in a quantitative yield. Amide coupling using HATU as a coupling reagent yielded the simple amides **4.33a** and **b**, after which, reductive amination yielded the final products **4.34a** and **b** (Scheme 4.34).



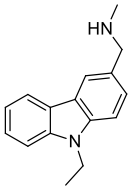
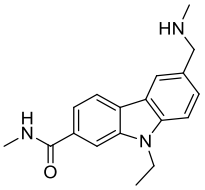
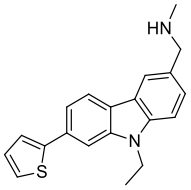
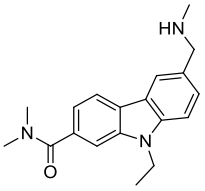
4.13 | **4.33,** 80% | R = H (**4.33a,** 90%), Me (**4.33b,** 84%) | R = H (**4.34a,** 27%), Me (**4.34b,** 14%)

Scheme 4.34: Development of conditions for an amide coupled library. (a) LiOH, MeOH, H_2O, 90 °C, 2 h; (b) HNMeR, HATU, NEt_3, DMF, rt; (c) $H_2NMe.HCl$, $Ti(O^iPr)_4$, NEt_3, DCM, EtOH rt 18 h then $NaBH_4$, rt, 8 h.

In both cases, for the SM library and amide library, only two simple final products were synthesised, as the feasibility of perusing each library of compounds was intended to be established, through SAR, before the synthesis of many diverse examples, i.e. is an amide linker tolerated in the narrow gap that bridges the central cavity and subsite II?

4.5.5 Initial Subsite II SAR

The synthesised products, from the SM library and amide library optimization, as well as all late stage intermediates, were assessed for their effect on the p53-Y220C DNA-binding domain using differential scanning fluorimetry as an initial assay. Any compounds that showed a marked improvement were reconfirmed, and their dissociation constants determined, using isothermal scanning calorimetry (Table 4.3).[§] Correlation between the two assays was good, showing a negative correlation of -0.84.

Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]	Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]
PK083		1.1	124	**PK9286 (4.34a)**		0.9	-
PK9318 (4.31a)		3.6	3.3 ± 0.3	**PK9287 (4.34b)**		0.2	-

[§] All biophysical assays were performed at the LMB by Dr Matthias Bauer.

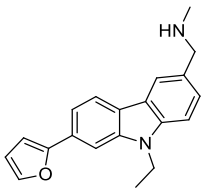
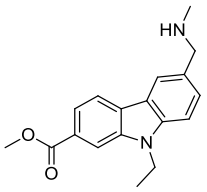
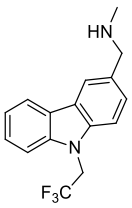
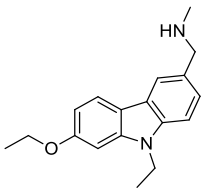
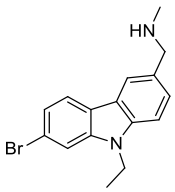
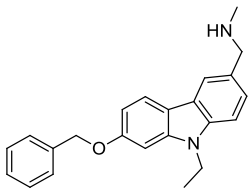
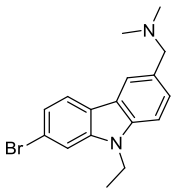
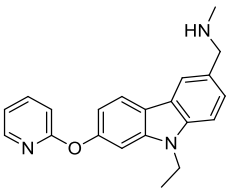
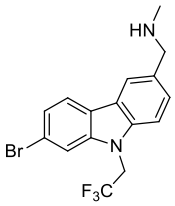
Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]	Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]
PK9320 (4.31b)		3.7	2.8 ± 0.7	**PK9288 (4.14)**		2.6	-
PK9255 (3.61)		1.4	35 ± 2.6	**PK9295 (4.35)**		1.3	62 ± 6.5
PK9284 (4.3a)		3.0	14.2 ± 2.0	**PK9296 (4.36)**		1.9	-
PK9285 (4.3b)		2.9	-	**PK9304 (4.28)**		-0.6	-
PK9301 (4.10)		3.7	5.9				

Table 4.3: Initial SAR of synthesised PK083 analogues that target subsite II. [a] All compounds were assayed as the amine·HCl salt to aid aqueous solubility. DSF and ITC assay conditions are described in Chapter 6. Correlation between DSF stabilization and ITC dissociation constant = -0.84.

Generally, most compounds summarized in Table 4.3 show increased stabilization of the p53-Y220C DNA binding domain over PK083, except in the case of the amides and pyridine ether. This is reasoned to be due to the increased steric bulk in direct proximity to the central carbazole ring clashing with the narrow gap that bridges the central cavity and subsite II.

The largest improvements in affinity occur when R^3 is heterocyclic or a bromide. Furthermore, the subsite II analogue of PK9255, PK9301, where R^3 is a bromide showed

an additive improvement in affinity. The binding modes of PK9284, PK9318 and PK9320 were determined by x-ray crystallography (Figure 4.11).**

For all three cases, there were two molecules in the asymmetric unit (corresponding to chains A and B). In all cases, electron density corresponding to the ligand was detected in both chains, where the ligand binding modes were identical. However, for PK083, electron density of the ligand was higher in chain B, hence, to draw comparisons with PK083, all numbers refer to chain B.

The binding mode of all three subsite II targeting carbazoles were almost identical to PK083, where the benzylic amine forms a hydrogen bond with the backbone C=O of Asp228. A slight shift of the carbazole in PK9318 and PK9320 compared to PK083 was observed in order to accommodate the 5-membered ring in subsite II (Figure 4.11).

Binding affinity can be dissected into two components, enthalpic and entropic. Contributions that are enthalpic arise from specific molecular interactions, whereas, entropic binding arises from nonspecific interactions, such as hydrophobic interactions and Van der Waals. In all three cases, the observed increase in binding affinity cannot be ascribed to any newly formed specific ligand-protein interaction so the increase in affinity is attributed to entropic gains. In the case of PK9284, substitution of the (ar)C-H for an aryl bromide can be viewed as affecting two properties, steric bulk and lipophilicity, where affinity gains arise from hydrophobic interactions with the lipophilic subsite II (See Section 3.1.4.1 for review of hydrophobic interactions).

The structures of PK9318 and PK9320 are virtually identical. Incorporation of the heterocyclic ring results in a shift of the Cys220 backbone carbonyl as well as a flip of Glu221 in order to accommodate the heteroaromatic moiety. For PK9318, an observed sulfur-sulfur contact of 3.7 Å occurs. Similarly to PK9284, increases in binding affinity can be attributed to hydrophobic interactions between the aromatic ring and the hydrophobic interaction surface of subsite II (Figure 4.11).

** X-ray crystallography was carried out by Dr Andreas Joerger at LMB.

However, close examination of the structures of PK9318 and PK9320 reveals that the heterocyclic ring sub-optimally occupies subsite II, hence, binding affinity may be increased by extension of the aromatic ring (Figure 4.11). Furthermore, the cLogP of these compounds is very high which, in combination with the most potent substitution at position R (trifluoroethyl), as accomplished in PK9301, results in a dramatic decrease in aqueous solubility (Table 4.4). Hence, further optimization of the heterocyclic ring should

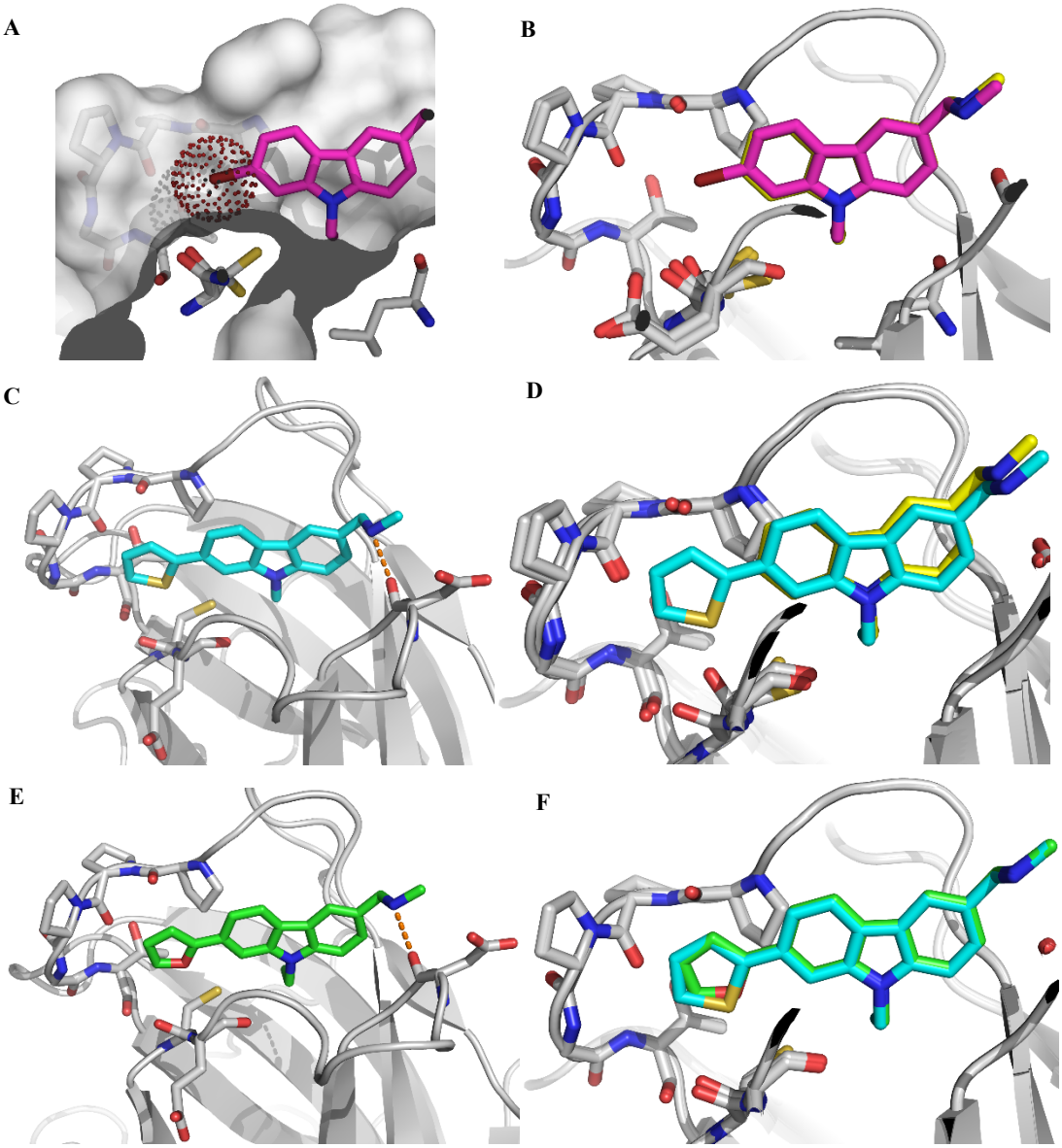


Figure 4.11: Co-crystal structure of PK9284 (A, pink), PK9318 (C, cyan) and PK9320 (E, green) bound to p53-Y220C determined at a resolution of 1.55, 1.24 and 1.32 Å respectively. (B) Comparison of PK9284 with PK083 (yellow), PK9318 with PK083 (D) and PK9318 and PK9320 (F). Hydrogen bonds are shown in orange dashed lines.

seek to (i) modulate the lipophilicity toward more soluble compounds and (ii) establish the steric limits by synthesis of bulkier heterocycles.

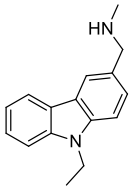
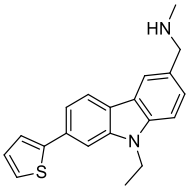
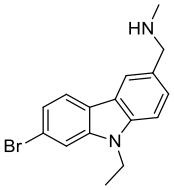
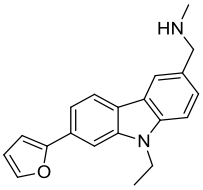
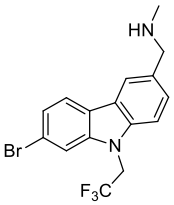
Compound ID	Structure	ITC K_d (µM)	cLogP	Compound ID	Structure	ITC K_d (µM)	cLogP
PK083		124	3.23	**PK9318**		3.3	4.65
PK9284		14.2	4.00	**PK9320**		2.8	3.94
PK9301		5.9	4.60				

Table 4.4: cLogP of highlighted subsite II targeting compounds that show higher potency toward p53-Y220C. Calculations were performed using logP prediction in MarvinSketch 14.8.25.0.

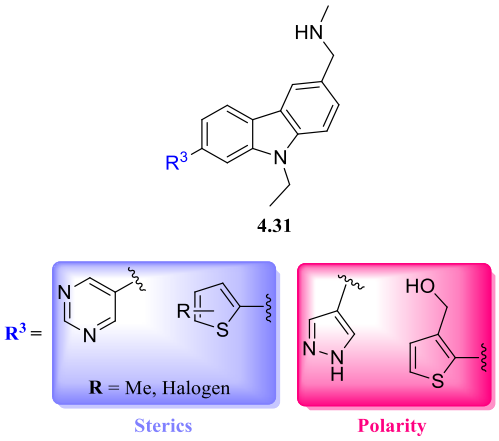


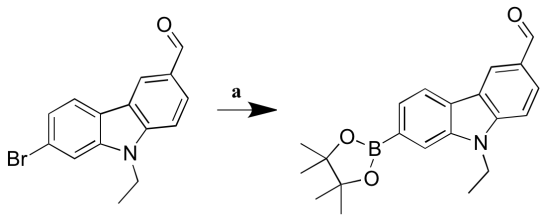
Figure 4.12: Strategies in optimization of the heterocyclic ring.

4.5.6 Optimization of Heterocyclic ring

Optimization of the heterocyclic ring, as outlined in Figure 4.12, was conducted firstly by establishing the most potent heterocyclic substitution and secondly, through exploration of the steric and electronic limits.

In order to expand the choice of heterocyclic coupling partners, (limited availability of heterocyclic boronic acids compared with heterocyclic bromides), the intermediate **4.37**

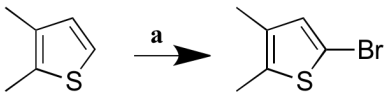
was synthesised from **4.4** using a Miyaura borylation reaction in an excellent yield (Scheme 4.35).



4.4 **4.37,** 95%

Scheme 4.35: Miyaura borylation of **4.4**. (a) $(BPin)_2$, $Pd(dppf)Cl_2.CH_2Cl_2$, KOAc, 1,4-dioxane, 80 °C, 1h

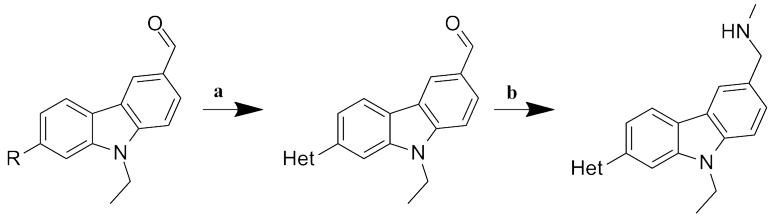
In the case that the desired heterocyclic coupling partner was not commercially available, e.g. 5-bromo-2,3-thiophene, bromination was carried out using the conditions summarized in Scheme 4.36.



4.38, > 99%

Scheme 4.36: Bromination of methyl-thiophenes. (a) NBS, DCM, rt, 18 h.

Using the optimized procedure outlined in Scheme 4.33, a number of different heterocyclic products were synthesised in varying yields (Table 4.5). In general, coupling of all unprotonated heterocycles (thiazole, furan, thiophene, isoxazole) proceeded in good yields. Coupling of protonated heterocycles, such as pyrazole, was unsuccessful and led to recovery of starting material. As outlined in Scheme 4.10, unprotected protonated heterocycles can undergo side reactions / cause inactivation of the catalyst. This problem can be circumvented by coupling the N-H protected analogues. Coupling of *N*-boc-pyrazole-4-boronic acid pinacol ester with intermediate **4.30h,** which upon reductive amination yielded the desired final product **4.31h** after treatment with HCl to both deprotect and synthesise the amine hydrochloride salt.



Scheme 4.37: Generic scheme for the synthesis of subsite II heterocycles. (a) Het-R^1, K_2CO_3, Pd-118, MeCN, H_2O, 90 °C; (b) H_2NMe, $NaBH(OAc)_2$, THF, rt.

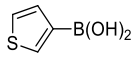
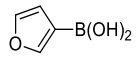
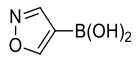
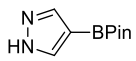
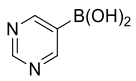
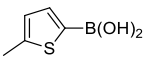
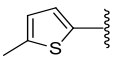
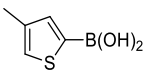
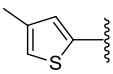
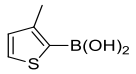
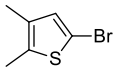
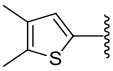
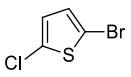
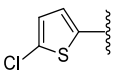
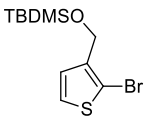
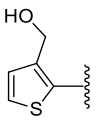
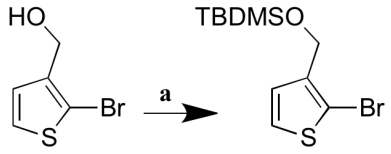
	Step A			Step B		
Compound	**R**	**Het-R[1]**	**Yield (%)**	**Compound**	**Het**	**Yield (%)**
4.30c	Br		74	**4.31c**		22
4.30d	Br		59	**4.31d**		61
4.30e	BPin		99	**4.31e**		48
4.30f	BPin		30	**4.31f**		10
4.30g	Br		49	**4.31g**		18
4.30h	Br		16	**4.31h**		29
4.30i	Br		46	**4.31i**		quant
4.30j	Br		38	**4.31j**		43
4.30k	Br		90	**4.31k**		28
4.30l	BPin		75	**4.31l**		39
4.30m	BPin		46	**4.31m**		47
4.30n	BPin		88	**4.31n**		14
4.30o	BPin		70	**4.31o**		19

Table 4.5: Subsite II heterocycle library synthesis.

In the case of **4.30o**, where the coupling partner contained a benzylic alcohol, the reaction led to a complicated mixture of products. Protection of the alcohol as a silyl ether limited side-reactivity and resulted in the formation of the desired product. Reductive amination followed by silyl ether deprotection yielded the final product **4.31o**.

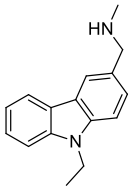
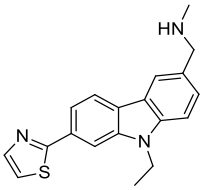
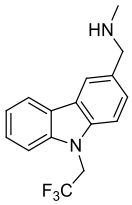
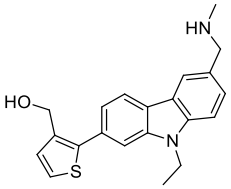
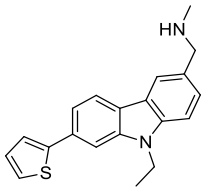
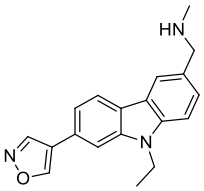
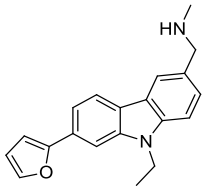
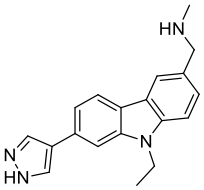
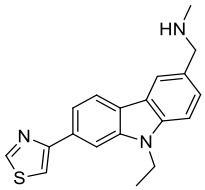
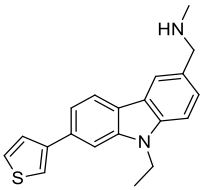
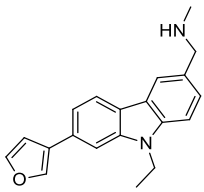
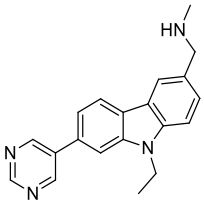


4.39, 53%

Scheme 4.38: Sily protection of a benyl alcohol functionalized thiophene. (a) TBDMSCl, Et_3N, DCM, 0 °C to rt.

4.5.7 SAR of Heterocyclic Optimization

Compounds **4.31c-o** were assayed, using the same standard procedure as described previously, against p53-Y220C using DSF as a primary screen and ITC to validate hits. Correlation between the two assays was good, showing a negative correlation of -0.85.††

Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]	Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]
PK083		1.1	124	**PK9322 (4.31f)**		3.0	8.3 ± 2.6
PK9255 (3.61)		1.4	35 ± 2.6	**PK9325 (4.31o)**		1.9	-
PK9318 (4.31a)		3.6	3.3 ± 0.3	**PK9324 (4.31g)**		1.8	-
PK9320 (4.31b)		3.7	2.8 ± 0.7	**PK9326 (4.31h)**		0.8	-
PK9323 (4.31e)		3.4	3.9 ± 0.4	**PK9319 (4.31c)**		0.5	-
PK9321 (4.31d)		3.2	5.6 ± 1.1	**PK9305 (4.31i)**		0.3	-

†† DSF and ITC data were determined by Dr Matthias Bauer.

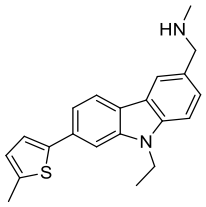
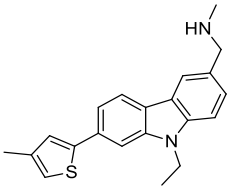
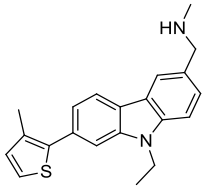
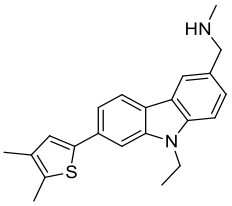
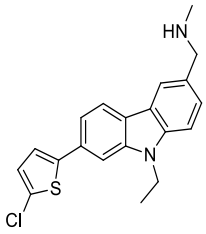
Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]	Compound ID	Structure[a]	DSF T_M [250 uM]	ITC K_D [uM]
PK9327 (4.31j)		3.1	-	**PK9328 (4.31l)**		3.3	1.8 ± 0.7
PK9331 (4.31k)		1.7	-	**PK9329 (4.31m)**		0.1	-
PK9332 (4.31l)		2.4	-				

Table 4.6: SAR of subsite II heterocyclic group scan. [a]Compounds were assayed as the amine.HCl salt to aid aqueous solubility. DSF and ITC assay conditions are described in Chapter 6. Correlation between DSF stabilization and ITC dissociation constant = -0.84.

The DSF and ITC data summarized in Table 4.6 reveal that, by varying the heterocyclic ring, no substantial improvements of PK9318 and K9320 were made. The regio-isomeric 3-furan and 3-thiophene substitutions both resulted in a reduction in affinity. Interestingly, the 3-thiophenic substitution resulted in loss of binding whereas 3-furan retained activity (5.6 μM). PK9323 and PK9322, which correspond to 4-thiazole and 2-thiazole substitutions respectively, are both potent binders (3.9 and 8.3 μM respectively), albeit slightly less potent than PK9318 and PK9320. Attempts to modulate the cLogP through introduction of more 'polar' heterocycles, for example, 4-pyrazole (PK9326) were unsuccessful, further reinforcing that lipophilic substitutions are most potent in subsite II. Furthermore, 6-membered heterocycles were not tolerated in the subsite II position, as exemplified by PK9305. Therefore, attempts to further increase affinity through functionalization of the heterocyclic ring were conducted using the 2-thiophenic building blocks owing to the fact that PK9318 is one of the most potent analogues and the desired building blocks were easy to obtain commercially or via a one-step bromination reaction (Scheme 4.36).

A chlorinated-thiophene in position-5 led to a reduction in affinity compared to 5-methyl (PK9327) indicating an upper lipophilicity threshold for subsite II substitutions. Introduction of a methyl group (PK9331) and benzylic alcohol (PK9325) at position 3- were generally not well tolerated, resulting in a reduction of stabilization by DSF, indicating that sterically-hindered substituentsat this position are not favoured.

Methylation of the thiophene ring in either the 4- or 5- position led to an improvement in affinity, for example, 4-methyl thiophene (PK9328) exhibited a dissociation constant of 1.8 μM. However, methylation of both positions 4,5- (PK9329) resulted in loss of binding.

The structures of PK9327 and PK9328 bound to p53-Y220C were solved in order to provide structural reasoning for the differences in potency between the di-methylated PK9329 and the mono-methylated PK9327 and PK9328 (Figure 4.13).‡‡ Comparison of binding modes of PK9327 and PK9328 with PK9318 reveals that the carbazole core adopts a virtually identical position in the central cavity between Val147, Pro223, Pro222 and Pro151. In both cases the benzylic amine forms a hydrogen bond between the N-H and the backbone carbonyl of Asp228 (Figure 4.13C, D). Comparison of a 4-methyl thiophene (PK9328) with thiophene (PK9318) substituent reveals that, in order to accommodate the methyl substituent, the thiophene ring rotates 15° relative to carbazole, away from the hydrophobic interaction surface formed by Pro152, Pro153, departing from the coplanar geometry exhibited in PK9318 (Figure 4.13A). Comparison of 5-methyl thiophene (PK9327) with PK9318 reveals that, similarly to PK9328, the thiophene ring rotates in order to accommodate the methyl substituent, however, this time toward the hydrophobic interaction surface causing the side-chain of Glu221 (E221) to flip (Figure 4.13B, F).

Examination of the structures of PK9327 and PK9328 reveals that 4,5-methyl substitution cannot be tolerated as it would cause a steric clash either with the backbone carbonyl of Cys220 or Pro152 (Figure 4.13E).

‡‡ The crystallography was carried out by Dr Andreas Joerger, LMB.

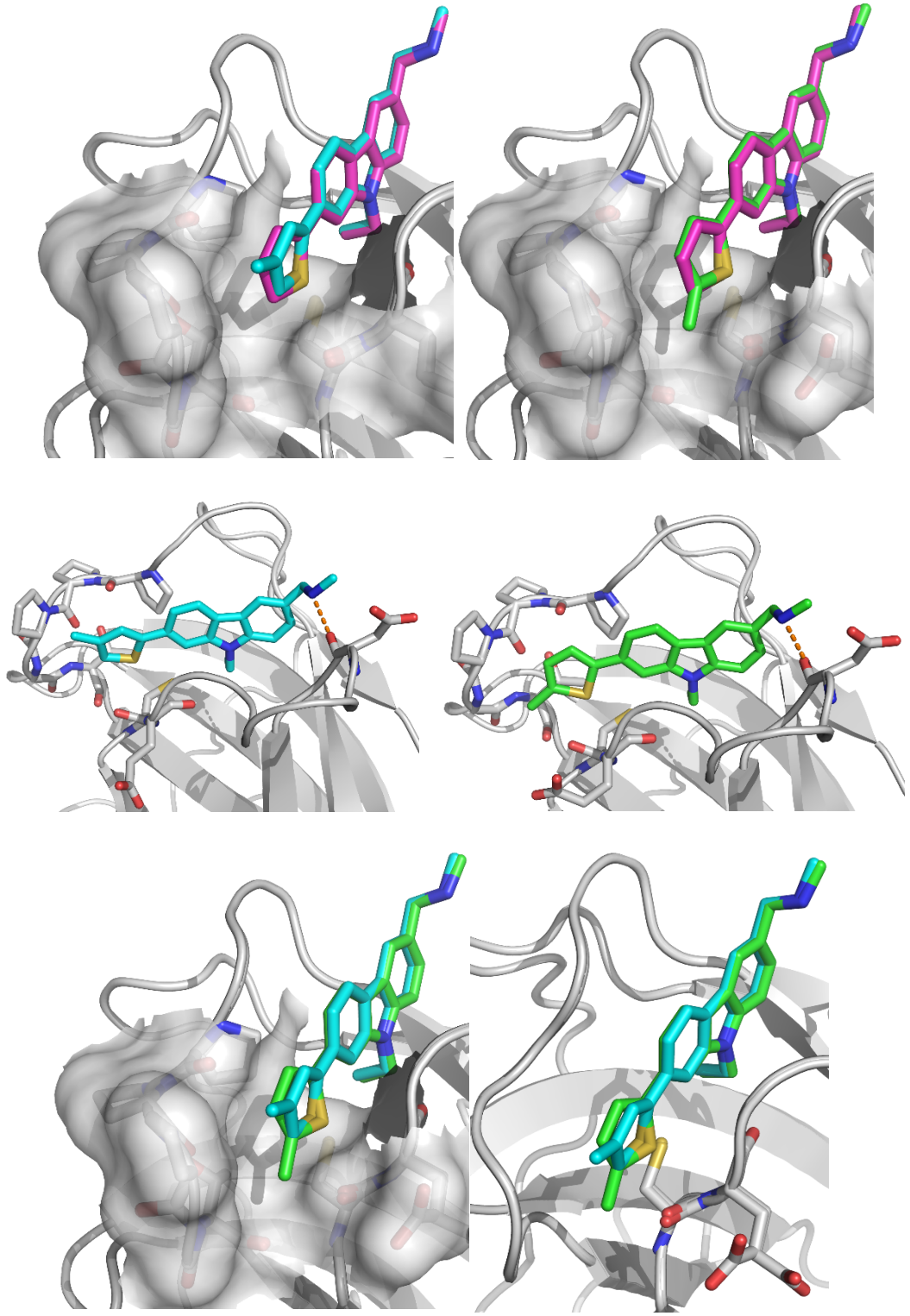


Figure 4.13: Structures of PK9327 and PK9328 bound to p53-Y220C. Structures of the binding modes of PK9328 (A, 1.32 Å resolution) and PK9327 (B, 1.25 Å resolution) compared to PK9318. Structures of PK9328 (C) and PK9327 (D) bound to p53-Y220C. Comparison of the binding modes of PK9327 and PK9328 (E) and highlighted induced E221 flip (F). Colour scheme: hydrogen bonds are represented as orange dashed lines.

4.5.8 Cell-Based Assays

Previously, identification of p53-activating compounds by using antibody and cell-based assays has resulted in false positives, where apparent p53 reactivation of mutants has later been attributed to the more general effects of cytotoxicity (such as DNA intercalators), or more specific effects that manifest themselves as p53 reactivation.

The results presented in this thesis so far present the optimization of PK083 analogues towards the p53-Y220C mutation-induced pocket, seeking to increase affinity using biophysical techniques such as ITC or $^{15}N/^{1}H$ HSQC as an assessment for compound potency. Whilst this, in combination with crystallography studies, confirms that PK083 analogues bind the mutation-induced pocket, they fail to establish whether the optimized compounds show Y220C specificity, manifesting as upregulation of p53 target genes in cell-based assays, or inhibit p53 aggregation. The utilization of these two techniques (biophysical and cell-based) in parallel helps to establish the most potent, selective compounds towards a lead compound optimized using a multi-parameter approach.

4.5.8.1 Cell Viability

Cell viability assays measure the ability of cells to maintain or recover viability where viability ranks on an index between 0 and 1 corresponding to cell life and death. Hence, cell viability can provide insight into compound toxicity. Furthermore, in the case of p53, by comparison of cell viability in cell lines that display a Y220C mutation and wt-p53 cell lines, the selectivity of compounds can be crudely established.

Figure 4.14 describes the viability of a wt-p53 cell line, NUGC-4 versus viability of a gastric adenocarcinoma cell line that carries a Y220C mutation, NUGC-3 upon treatment with PK9318, PK9320, PK9323 and PK9328.[§§] Generally, all compounds appear to cause cell death at 20 μM in a non-specific manner, indicating compound toxicity. This pattern is observed at a ligand concentration of 10 μM for PK9318, PK9320 and PK9328. However, PK9323 appears to show some Y220C selectivity, with almost complete cell death of Y220C cells vs 40% viability of the wild-type associated cell-line. At 3 μM, this

[§§] Cell viability assays were carried out by Dr Matthias Bauer.

Y220C-specific effect is further exemplified, with p53-Y220C cell-lines retaining 40% viability versus 85% viability for the wild-type associated cell-line.

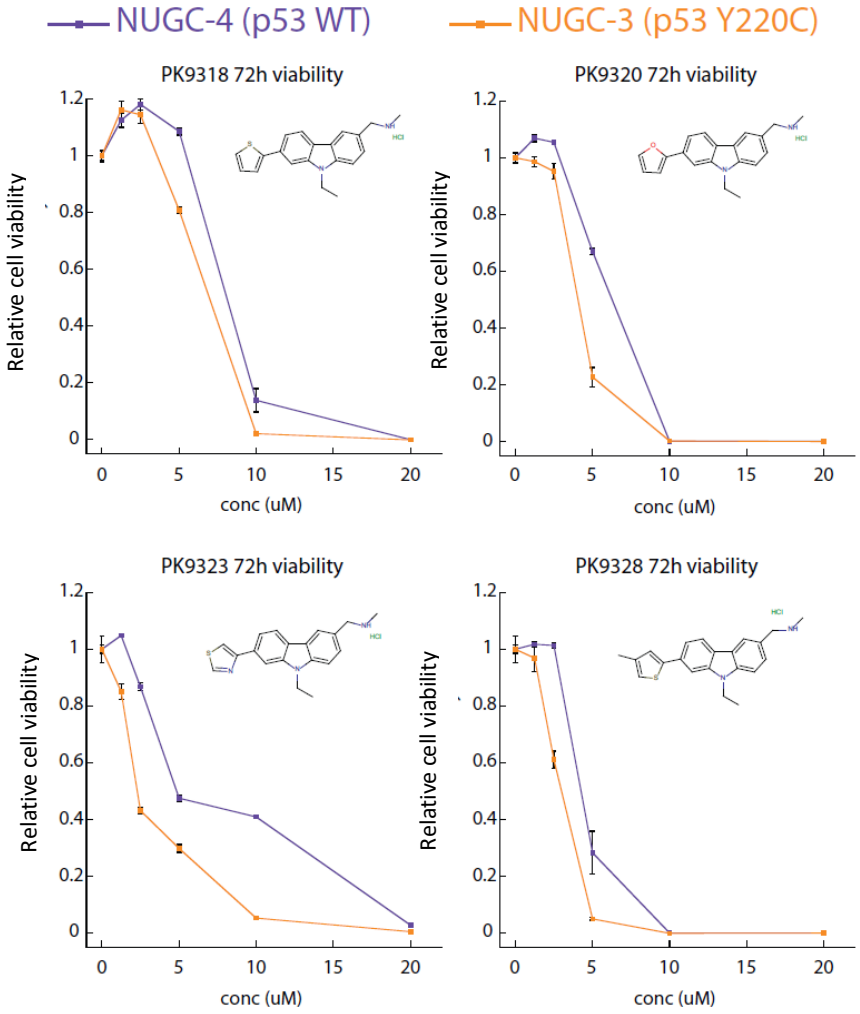
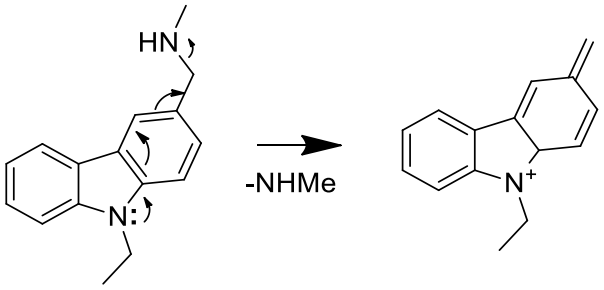


Figure 4.14: Cell Viability of PK9318, PK9320, PK9323 and PK9328.

In summary, the cell viability assays indicate that the subsite II heterocyclic compounds all show behaviour indicative of inherent non-specific toxicity at higher concentrations, with only PK9323 exhibiting Y220C specific effects at lower concentrations. Speculatively, the toxicity of the carbazole series may be due to:

i) Genotoxicity related to DNA-intercalation owing to the planar structure of PK083 and analogues. N.B. Carbazoles constitute a commonly used motif for the design of cytotoxic agents.[525]

ii) Formation of reactive decomposition products arising from loss of benzylic amine that is assisted by the carbazole lone pair. Commonly observed by mass spectrometry (Scheme 4.39).



Scheme 4.39: Postulated formation of a reactive decomposition product of PK083.

4.5.8.2 Initial Compound PK Properties

Additionally to cell viability, some pharmokinetic properties of PK9318, PK9320 and PK9323 were determined (Table 4.7).[***] Assays such as microsomal stability, plasma protein binding, and HepG2 cytotoxicity, aimed at assessing and predicting PK083 analogues metabolic liability, *in vivo* hepatic clearance, bioavailability, preponderance towards hERG inhibition, and toxicity were conducted to provide results with which to inform future compound design. Below, desirable PK083 analogue ADME and toxicity parameters are outlined.

i) *Microsomal turnover* (human/rat) to assess metabolic liability, to predict *in vivo* hepatic clearance, looking for low intrinsic clearance values (<2 μg/min/mg protein).

ii) *Plasma protein binding* to assess unbound (bioavailable) drug: desirable <95%; undesirable high binding (>95%). High plasma protein binding is indicative of low bioavailability (lower concentration of drug in bloodstream as bound to plasma proteins) and a high preponderance towards hERG/toxic effects (via inhibition of the hERG K^+channel, where SAR indicates hydrophobic or cationic moieties as toxicophores within drugs).[526]

iii) *HepG2 Cytotoxicity* to assess off-target toxicity. Desirable EC_{50} (HepG2) > > EC_{50} (p53-Y220C) to ensure compounds has a therapeutic rather than toxic effect.

[***] Assays were carried out at Selcia Ltd.,by Dr Martin Walker

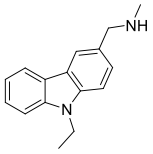
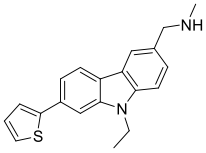
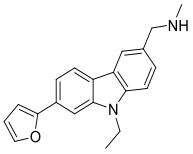
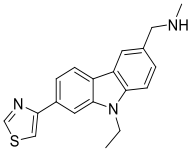
Compound	Structure	Microsome Stability[a]		Plasma Protein Binding[b]		HepG2 Cytotoxicity	
		$T_{1/2}$ (min)	Intrinsic Clearance (µg/min/mg protein)	% unbound	% Bound	EC_{50} (µM)	% inhibition at top concentration[c]
PK083		45.6	30	9.0	91.0	23.0	100.0
PK9318		61.6	22	0.0	100.0	3.5	99.9
PK9320		24.8	56	0.2	99.8	4.8	100.1
PK9323		74.9	18.5	3.0	97.0	1.1	100.0

Table 4.7: Pharmacokinetic properites of PK083, PK9318, PK9320 and PK9323. [a]The species used was mouse. The assay contained 0.20% DMSO and 5 µM compound f/c. [b]The species used was mouse. The assay contained 10 µM compound and 0.1% DMSO f/c. [c]Top concentration of 100 µM. Colour scheme: yellow, medium; red, high.

The microsomal stability assay assesses metabolic stability using microsomes obtained by homogenization of a liver followed by centrifugation of the homogenate to yield a supernatant fraction (S9 or S10). A microsomal pellet is then prepared by centrifugation of the S9 or S10 fraction and contains the smooth endoplasmic reticulum where the enzymes responsible for phase I oxidation, including the cytochrome P450 monooxygenases, reside. With the exception of uridine 5'-diphospho-glucuronosyltransferase (UGT), most phase II enzymes are cytosolic are, therefore, absent from liver microsomes. Microsomes are supplemented with the cofactor NADPH. With the exception of PK9320, the subsite II substitution generally increases microsomal stability from a half-life of 45 minutes for PK083 to 61 minutes and 74 minutes for PK9318 and PK9323 respectively. In all most cases, except for PK9323, degradation of compound was related to CYP450 metabolism, however, degradation of PK9323 in the

NADPH sample was similar to the +NADPH sample, suggesting non-CYP450 metabolism (Figure 4.15).

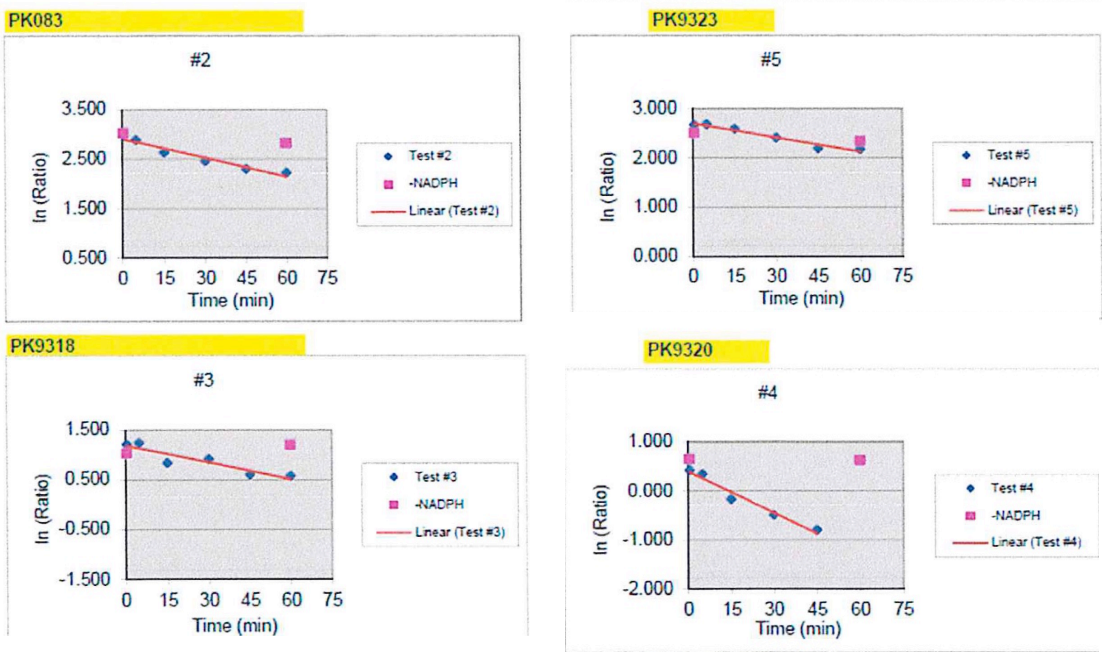


Figure 4.15 Mouse microsomal stability.

Another helpful parameter that can be determined that aids in estimating oral bioavailability is plasma protein binding. Human serum albumin (HAS) is the most prominent protein in plasma and binds different classes of ligands at different sites, for example, bulky heterocyclic anions bind to Sudlow's site I, whereas, aromatic carboxylates Sudlow's site II.[527] HSA provides a depot for many compounds, where the extent to which compounds are bound can adversely affect the pharmokinetics and pharmodynamics of drugs. The subsite II modifications of PK9318, PK9320 and PK9323 adversely affect plasma protein binding compared to PK083, with 91% bound versus 100, 99.8 and 97% respectively.

Inhibition of HepG2 is commonly used to identify cytotoxic compounds. Results of the HepG2 inhibition assay further reinforce the observed off-target toxicity observed in the cell viability data, where all compounds showed enhanced toxicity over PK083.

4.5.8.3 H2AX Phosphorylation Assay

In order to further elucidate the potential underlying mechanism/s behind the observed non-specific off-target effects that the compounds exhibited in the cell viability assay, as

well as the observed cytotoxicity identified by the HepG2 assay, a histone H2AX phosphorylation assay was conducted. In response to ionizing radiation or DNA-damaging chemotherapeutic agents, double-stranded breaks (DSBs) are generated that rapidly result in the phosphorylation of histone H2A variant H2AX. Because phosphorylation of H2AX at Ser139 (γ-H2AX) is abundant, fast and correlates with DSBs, it is the most sensitive marker for identification of DNA damage and the subsequent repair. Therefore, it is commonly used to identify compounds that are genotoxic. [528]

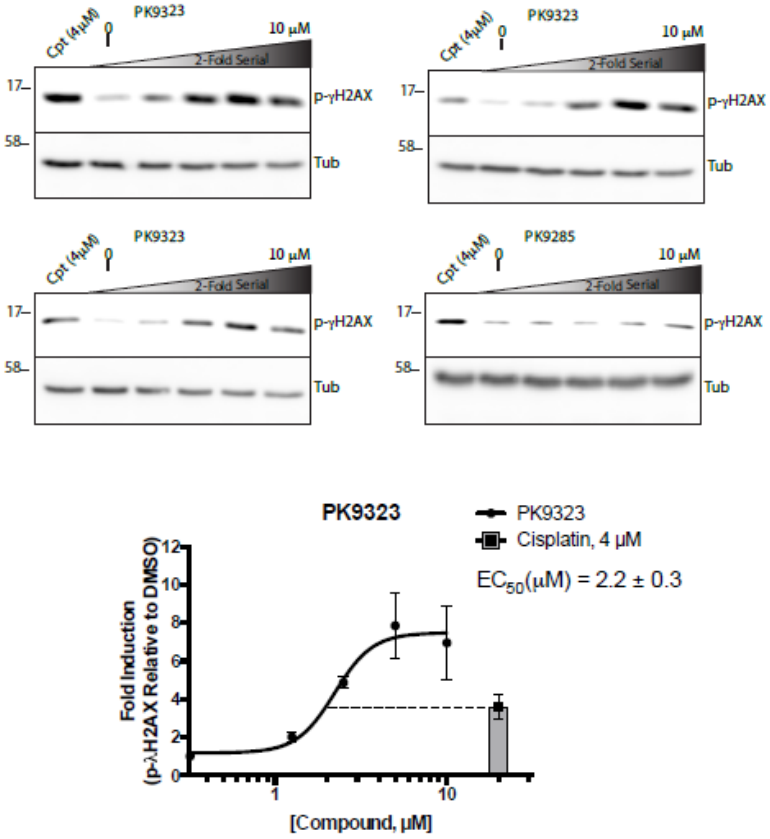


Figure 4.16: H2AX phosphorylation assay. Experiment. Cells (MCF7 breast cancer) plated in triplicate in 12-well dish – 100,000 cells/well. Cells were allowed to settle for 2d, then treated with the indicated concentration of compound. After 20 h, the cells were lysed with 1X SDS-loading buffer, heated to 95 °C for 5 min, sonicated, clarified by centrifugation (13 KRPM, 10 min), and proteins separated by SDS-PAGE. Proteins were then transferred to PVDF membrane and blotted with the indicated mouse derived antibody. Goat derived, anti-mouse HRP-conjugated antibody (DAKO) was used as a secondary antibody for luminescent detection of proteins using ImageQuant.

Compounds PK083, PK9320 and PK9323 were tested for their genotoxicity using the H2AX phosphorylation described in .[†††]PK083 and PK9320 do not induce DNA damage,

[†††] Assay was carried out by Dr Cory Ocasio.

however, PK9320 showed a great deal of cytotoxicity as Tubulin levels decrease dramatically, indicating cell death or inhibition of cell growth (unpublished results, Tony). Paradoxically, PK9323 was identified as highly genotoxic, with an EC_{50} of 2.2 μM, out-performing Cisplatin (EC_{50} = 4 μM). By comparison with cell viability results that show PK9323 had a Y220C specific effect at 10 μM, which is associated with a reduction in off-target toxicity compared the other compounds, the H2AX phosphorylation assay show it is the most genotoxic. Efforts are ongoing to attempt to rationalize this observation by elucidation of the toxicity mechanism(s) to inform subsequent compound design.

4.6 Conclusions

Targeting subsite II for affinity gains through functionalization of R^3 of the generic carbazole based on PK083 represented in Figure 4.1 proved successful. Synthesis of a 2-brominated PK083 analogue, **4.44**, via a two-step synthesis, provided a rapid synthesis toward a late-stage intermediate for funcitonalization, where reductive amination of the resulting diversified library provided a nice handle for purification of the final products. However, application of this synthetic strategy to the synthesis of the late-stage intermediate **4.10**, which is an analogue of PK9255, proved unsuccessful, where the powerful inductive electron-withdrawing nature of the *N*-3,3,3-trifluoroethyl group caused the intermediate **4.7** to be unreactive towards Vilsmeier-Haack formylation conditions. Synthesis of **4.7** was accomplished through a 3-step synthesis, based on reductive cyclization of the nitrobiphenyl **4.9**, in which the aldehyde functionality was already installed.

Biophysical characterization of an initial library of subsite II carbazoles, using DSF as a primary screen and ITC to validate hits, identified the compounds PK9318 and PK9320, synthesised by Suzuki-Miyaura reaction of intermediate **4.4** with 2-thiophene boronic acid and 2-furan-boronic acid respectively, that showed enhanced stabilization of p53-Y220C (3.6 K and 3.7 K at 250 μM respectively) and Y220C binding affinity (K_D = 3.3 and 2.8 μM). Structural characterization of the Y220C binding modes revealed that the observed enhance binding affinity cannot be attributed to specific protein-ligand interactions and hence enhancements are attributed to favourable entropy arising from 'the hydrophobic effect.' This initial screen also identified the compounds PK9284 and

PK9301, which constitute the 2-brominated analogues of PK083 and PK9255, which showed enhanced p53-Y220C stabilization and binding affinity. Similarly to PK9318 and PK9320, enhancements were attributed to entropic effects via structural characterization of the Y220C binding mode of these compounds.

Furthermore, through library synthesis and the resulting SAR, the most potent heterocyclic substitution was identified as the original hits, where introduction of more 'polar' heterocycles as well as 6-membered rings were not tolerated. Methylation of the thiophene ring in positions 5- and 4- yielded the compounds PK9327 and PK9328 that showed enhanced stabilization and binding affinity, with DSF T_m = 3.1 (PK9327), 3.3 (PK9328) K. The dissociation constant was determined as 1.8 μM (ITC) for PK9328.

Whilst these compounds show enhanced stabilization and binding affinity of p53-Y220C through DSF and ITC, validation of enhanced p53-Y220C stabilization and Y220C specificity is still to be decisively established using cellular assays. Cell viability of compounds PK9318, PK9320, PK9323 and PK9328 was examined in an Y220C cell line vs a wild-type p53 cell line. Generally, at high concentrations (20 μM) these compounds reduce cell viability to 0 in both Y220C and wt-p53 cell-lines. However, at a lower compound concentration of 3 μM, both PK9323 and PK9328 reduce cell viability in the Y220C associated cells over the wild-type, indicating a Y220C specific effect. However, through HepG2 cytotoxicity and H2AX phosphorylation assays, PK9323 was identified as both highly cytotoxic and genotoxic, where PK9323 was determined to have an EC_{50} of 2.2 μM in the H2AX phosphorylation assay (MCF7 breast cancer cell line).

Future efforts to further optimize the PK083 fragment to p53-Y220C will concentrate on a multi-parameter optimization approach, in which affinity gains will be considered equally against cell viability as well as compound PK and toxicity. Strategies toward this are outlined in Chapter 5.

Chapter 5 : Conclusions and Future Directions

The aims of this thesis were to develop chemistry towards compounds based on the hit compound PK083 that show enhanced potency and selectivity towards the thermally unstable oncogenic p53 mutant Y220C, using SAR and x-ray crystallography to inform compound design. Ultimately, the aim of this project is to develop a small molecule p53-Y220C stabilizing drug, based on PK083 or other identified hits, as a personalized medicine approach to cancer therapy.

Examination of the structure of PK083 bound to p53-Y220C revealed that the PK083 binding site could be subdivided into 4 distinct sub-pockets (Figure 5.1), the central cavity (cyan) and subsites, I (red), II (green) and III (pink), which could be targeted through different strategies.

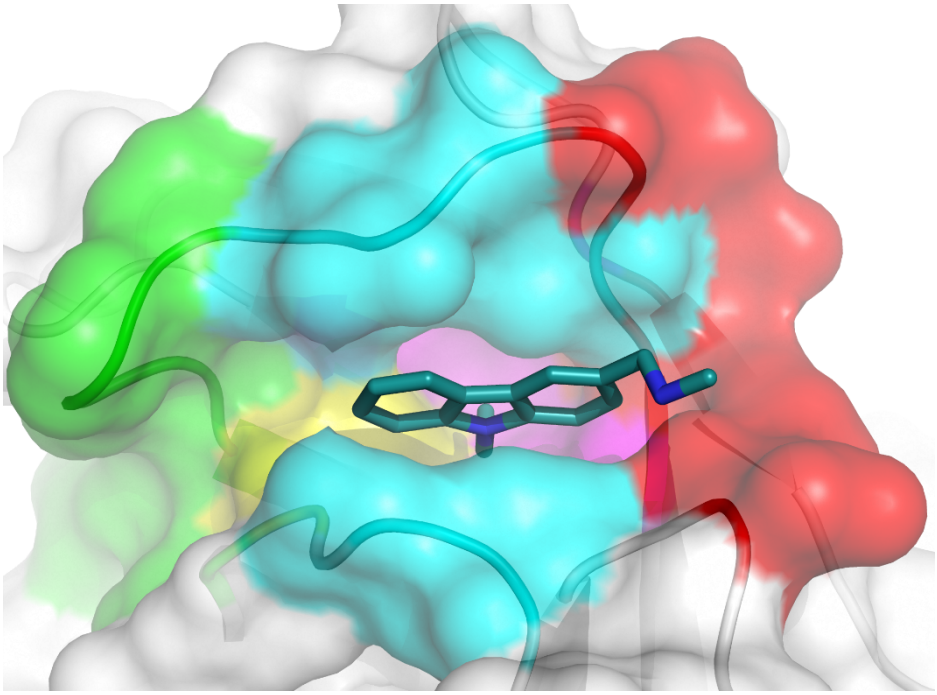


Figure 5.1: PK083 bound to p53-Y220C (PDB: 2VUK). Mutation induced cavity is subdivided into: Subsite I (red), Subsite II (green), central cavity (cyan), Y220C mutation (yellow) and subsite III (pink).

The solvent exposed subsite I had previously been targeted for affinity gains through targeting Asp228 (hydrogen bonding) and Trp1146 (π-π interactions) through diversification of the benzylic amine moiety. Despite a large amount of compounds purchased or synthesised, no significant improvements were made. Accordingly, other areas of the PK083 binding site were targeted for affinity improvements.

Chapter's 2 and 3 describe the synthesis of compounds to target the mutant specific, hydrophobic Cys220 sub-pocket through functionalization of R and R^1, which were accessed through alkylation chemistry (R) or reductive cyclization of the appropriate nitrobiaryl (R^1) (Figure 5.1, yellow).

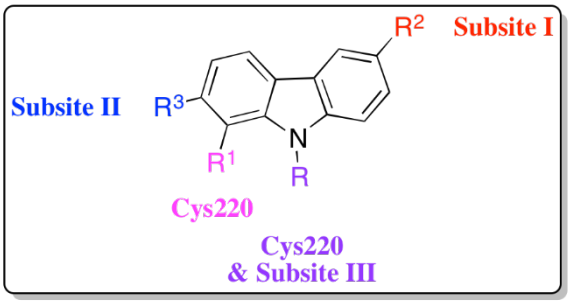


Figure 5.2: Structure of a generic carbazole with key growth vectors highlighted.

The steric limit of the mutational sub-pocket was defined through reductive amination of commercially available *N*-functionalized carbazole aldehydes. The SAR defined that R< *n*Pr or *i*Pr, providing a limit to inform compound design.

Attempts to target Cys220 through halogenation of R^1 via S···Cl or fluorine multipolar/hydrogen bonding were unsuccessful and led to a reduction in Y220C stabilization. However, incorporation of fluorine into the *N*-ethyl anchor at R proved more successful, yielding two compounds, PK9255 and MB065, with substantially increased p53-Y220C stabilization (1.2 K at 125 µM) and Y220C binding affinity. Compared with the parent compounds PK083 and PK8017, the ITC K_D values improved by a factor of 5. The structures of PK9255 and MB065 bound to Y220C confirmed that the binding mode of the carbazoles was almost identical to that of PK083, however, variations at R^2 caused the benzylic amine to flip, breaking the hydrogen bond with the backbone Asp228 C=O, forming a new hydrogen bond with the C=O of Val147. This was reasoned to be responsible for affinity differences between mono- and dimethylated PK083 analogues synthesised in this thesis. The CF_3 group interacts with the backbone carbonyl groups of Leu145 and Trp146 as well as with the thiol group of Cys220. Analysis of C–S···F angles in the complexes with di- and trifluorinated compounds suggests that fluorine interacts with Cys220 via weak hydrogen bonding with the polarized proton of the thiol function and via sulfur σ-hole bonding at an angle close to 180 °.

Following unsuccessful attempts to enhance affinity through targeting the transiently open subsite III by arylation of R via Ullmann coupling, subsite II (Figure 5.1, green) was targeted for affinity improvements via functionalization of R^3 (Figure 5.2).

Synthesis of privileged intermediates based on PK083 and PK9255, brominated at R^3 (**4.4** and **4.10)**, was accomplished via functionalization of 2-bromocarbazole (PK083) or reductive cyclization of the appropriate nitrobiphenyl (PK9255), and provided reactive intermediates to target subsite II. Additionally, the available chemical space for targeting subsite II was expanded by synthesis of other PK083 analogues, using the same method as for the synthesis of **4.10**, in which R^3 was a methyl ester (amide library). Attempts to synthesise a PK083 analogue where R^3 = OH, were unsuccessful. However, early-stage diversification of the nitrobiphenyl, for example, by Ullmann coupling, yielded a diversified product that followed by reductive cyclization, ethylation and reductive amination yielded an Ullmann coupled PK083 analogue. Conditions were sought for diversification of **4.4** (SM reaction) and **4.13** (amide coupling) and a few examples in each class synthesised.

Biophysical characterization of an initial library of subsite II carbazoles, comprised of R^3 functionalized PK083 late stage intermediates as well as initial SM (**4.4**) and amide coupled (**4.13**) analogues identified compounds PK9318 and PK9320, in which R^3 is heterocyclic, as the most potent substitution. PK9318 and PK9320 both showed enhanced stabilization of Y220C (3.6 K and 3.7 K at 250 μM) and Y220C binding affinity (K_D = 3.3 and 2.8 μM respectively). Structural characterization of the Y220C binding mode of these compounds confirmed the binding mode was almost identical to PK083, with both compounds forming a hydrogen bond between the backbone C=O of Asp228 and the NH of the benzylic amine. A slight shift of the carbazole ring is observed in order to accommodate the heterocyclic substitutions is subsite II. Introduction of the heterocycle does not result in the formation of specific protein-ligand interactions and hence, the observed enhanced affinity is attributed to entropic gains.

Optimization of the heterocyclic ring was accomplished by firstly, establishing the most potent heterocyclic substitution and secondly, through exploration of the steric and electronic limits of functionalization of the heterocyclic ring. Variation of the

heterocyclic ring revealed that the original hits, PK9318 and PK9320 were the most potent, with a 4-thiazole analogue (PK9323) exhibiting comparable Y220C stabilization (3.4 K at 250 μM) and dissociation constant (ITC K_D = 5.6 μM). Methylation of the thiophene ring, (PK9318) seeking to identify the steric limits of functionalization of the heterocyclic ring, revealed that positions 5 and 4 were both tolerated, where the 4-isomer (PK9328) showed enhanced binding affinity with an ITC determined K_D = 1.8 μM.

Whilst PK083 binding affinity has been successfully enhanced, conclusive validation that these more potent PK083 analogues stabilize p53-Y220C and show specificity (activation of p53 target genes and cell death, also inhibit aggregation) in cellular assays is yet to be determined. Generally, at high compound concentrations (20 μM), these compounds reduce cell viability to 0 in both Y220C and wild-type p53 cell lines, which is indicative of non-specific toxicity. However, at lower compound concentrations (3 μM), both PK9323 and PK9328 exhibit a Y220C specific effect, reducing cell viability in cell lines with Y220C p53 status over wt-p53 cell lines. Assessment of toxicity, using a HepG2 cytotoxicity assay, indicated that all subsite II targeting compounds were, to some extent, cytotoxic. Furthermore, a histone H2AX phosphorylation assay, which is commonly used to identify genotoxic compounds, identified PK9323 as highly genotoxic, inducing DNA damage in a breast cancer cell line (MCF7, N.B. MCF7 has wt-p53 status) with and EC_{50} of 4.4 μM. These results challenge the findings of the cell viability assay, paradoxically suggesting that PK9323 is highly toxic.

Studies are ongoing to attempt to elucidate the mechanism(s) behind this toxicity. Speculatively, the observed toxicity may be attributed to:

i) DNA intercalation owing to the planarity of the carbazole system (N.B. carbazoles constitute a commonly used compound class with which to design cytotoxic agents).[529]

ii) Formation of reactive decomposition products (e.g. that may arise from loss of the benzylic amine) forming protein-adducts leading to off-target toxicity (Scheme 4.39).

Future compound design should seek to:

i) Identify the mechanism behind the series toxicity.

ii) Increase compound aqueous solubility by modulating the cLogP to allow incorporation of the most potent subsite II substitution with the most potent mutational sub-pocket substitution.

iii) Further improve compound potency and selectivity, as assessed both through biophysical and cellular techniques.

Figure 5.3 demonstrates some strategies that could be used in order to improve PK083 analogue potency and selectivity. The genotoxicity of PK083 analogues has already been assessed by histone H2AX phosphorylation assay. As all compounds were not identified as genotoxic, non-specific toxicity cannot be solely attributed to the extended planarity of PK083 analogues causing genotoxicity through DNA intercalation. Alternatively, these analogues cause toxicity through formation of reactive decomposition products via loss of the benzylic amine leading to a Michael acceptor. By synthesis of the compound highlighted in Figure 5.3 in the yellow box, assessment of the contribution of the benzylic amine to the toxicity observed for PK083 analogues may be accomplished.

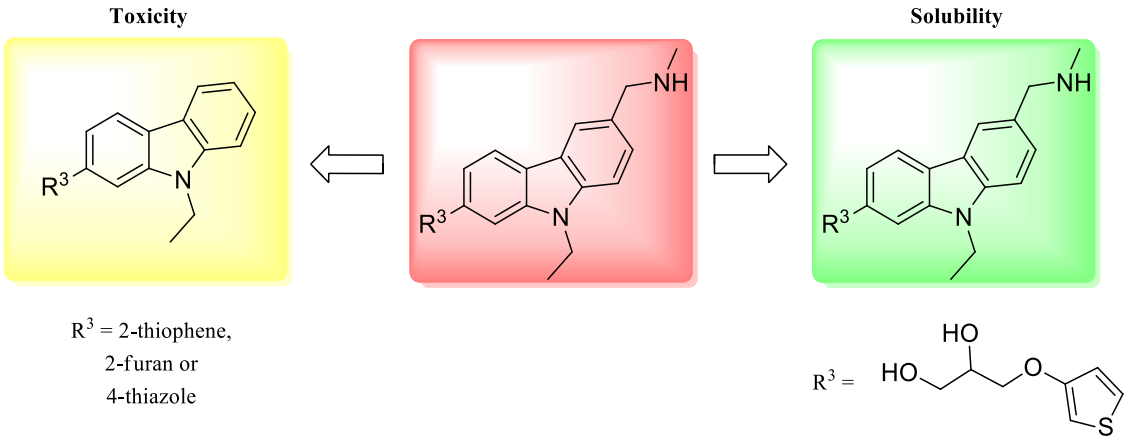


Figure 5.3: Summary of simple strategies to improve aqueous solubility and probe the mechanism behind the non-specific toxicity that PK083 analogues exhibited in cellular assays.

Chapter 6 :Experimental

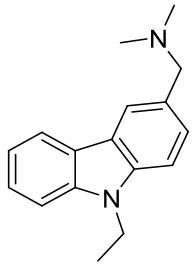
All reactions were carried out in air unless otherwise stated, using commercial grade starting materials, solvents and reagents. The progress of all reactions was monitored by TLC using commercially available glass silica gel plates (60 Å, F254). The mobile phase was generally a solvent mixture and the visualisation was undertaken using UV light. Microwave reactions were conducted in a CEM discovery microwave reactor. All NMR spectra were measured on either a Varian NMR 600, 500, or 400 MHz spectrometer. Chemical shifts are quoted in parts per million (ppm) (δ relative to a residual solvent peak for ^{1}H and ^{13}C). Chromatographic purifications were undertaken using an ISCO purification unit, Combi Flash RF 75 PSI, using Biotage silica gel columns. LC-MS purity analyses were undertaken using a 5 μm C18 110 Å column. Percentage purities were performed using a 30 min method in water/acetonitrile with 0.1% formic acid (5 min at 5%, 5–95% over 20 min, 5 min at 95% or 5 minutes at 30%, 30-90% over 20 minutes, 5 min at 90%) with the UV set to 254 nm unless otherwise indicated. All mass spectrometry was carried out at either the EPSRC UK National Mass Spectrometry Facility (NMSF) using a Thermo Scientific LTQ Orbitrap XL spectrometer or by Dr. Alaa Abdul-Sada at the University of Sussex. Caclulations of compound m/z for MS-EI were carried out using ChemDraw Ultra 12.0.3.1216.

6.1 Chapter 2

General Procedure A

Representative procedure for reductive amination using titanium (IV) isopropoxide as a reagent

Synthesis of 1-(9-Ethyl-9*H*-carbazol-3-yl)-N,N-dimethylmethanamine (**2.23**).



To a solution of 9-ethyl-9*H*-carbazole-3-carbaldehyde (98 mg, 0.44 mmol) in anhydrous EtOH/DCM (10 mL) was added dimethylamine hydrochloride (71 mg, 0.88 mmol), triethylamine (91 μL, 0.66 mmol), and titanium(IV) isopropoxide (260 μL, 0.88 mmol). The resulting solution was stirred at RT for 18 h before the addition of sodium

borohydride (33 mg, 0.88 mmol). The solution was stirred at RT for 8 h before pouring into 2 M aqueous ammonia (25 mL). The suspension was filtered through Celite, and to the filtrate was added H_2O. Crude product was extracted with DCM, dried over anhydrous K_2CO_3. The suspension was filtered, and the filtrate was concentrated in *vacuo* to yield crude product that was purified by chromatography on silica gel using DCM/MeOH 9:1 as an eluent to yield the product as a white solid, yield 82% (91 mg, 0.36 mmol).

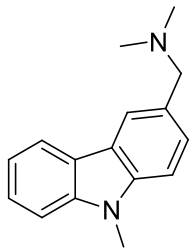
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.14−8.10 (m, 1H, ar CH), 8.06 (s, 1H, ar CH), 7.51− 7.34 (m, 4H, ar CH), 7.26−7.20 (m, 1H, ar CH), 4.37 (q, J = 7.2 Hz, 2H, CH_2), 3.64 (s, 2H, CH_2), 2.33 (s, 6H, CH_3), 1.45 (t, J = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 140.2, 139.3, 129.3, 127.1, 125.5, 122.9 (2C), 121.0, 120.4, 118.7, 108.4, 108.1, 64.7, 45.3 (2C), 37.5, 13.8 ppm.

HRMS-ESI (m/z) found 208.1120 $[M]^+$, calcd 208.1121 for $[C_{15}H_{14}N]^+$ (loss of NMe_2).

LC-MS purity 97% (UV), ret. time = 11.36 min.

Synthesis of 1-(9-Methyl-9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**2.24**).



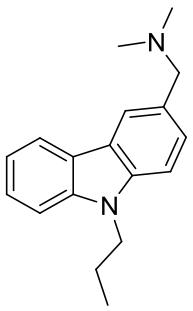
Synthesised according to general procedure A. 9-Methyl-*9H*-carbazole-3-carbaldehyde (305 mg, 1.46 mmol), dimethylamine hydrochloride (2 equiv.; 246 mg, 3 mmol, 2 equiv.), triethylamine (303 mg, 3 mmol), titanium isopropoxide (850 mg, 3 mmol) and sodium borohydride (85 mg, 2.3 mmol). Product isolated as a beige solid, yield 20% (68 mg, 0.28 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.11 (d, *J* = 7.8 Hz, 1H, ar CH), 8.05 (s, 1H, ar CH), 7.51 – 7.42 (m, 3H, ar CH), 7.40 (d, *J* = 8.2 Hz, 1H, ar CH), 7.36 (d, *J* = 8.2 Hz, 1H, ar CH), 7.24 (t, *J* = 7.4 Hz, 1H, ar CH), 3.86 (s, 3H, CH_3), 3.64 (s, 2H, CH_2), 2.32 (s, 6H, CH_3).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 141.2, 140.4, 127.1, 125.6, 122.7, 122.6, 120.9, 120.3, 118.8, 108.3, 108.0, 64.6, 45.2 (2C), 29.1 ppm (missing one quarternay carbon).

HRMS-ESI (m/z) found 194.0964 $[M]^+$, calcd 194.0964 for $[C_{14}H_{12}N]^+$ (loss of NMe_2).

Synthesis of 1-(9-Propyl-9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**2.25**)



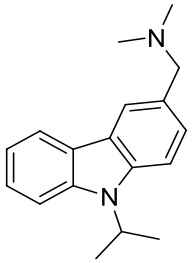
Synthesised according to general procedure A. 9-Propyl-*9H*-carbazole-3-carbaldehyde (980 mg, 4.13 mmol), dimethylamine hydrochloride (673 mg, 8.26 mmol), triethylamine (626 mg, 863 µL, 6.19 mmol), titanium isopropoxide (2.347 g, 2.445 µL, 8.26 mmol) and sodium borohydride (312 mg, 8.26 mmol). Product isolated as a brown oil, yield 27% (298 mg, 1.11 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.11 (d, *J* = 7.8 Hz, 1H, ar CH), 8.05 (s, 1H, ar CH), 7.49 – 7.43 (m, 3H, ar CH), 7.41 (d, *J* = 8.2 Hz, 1H, ar CH), 7.38 (d, *J* = 8.3 Hz, 1H, ar CH), 7.25 – 7.19 (m, 1H, ar CH), 4.28 (t, *J* = 7.2 Hz, 2H, CH_2), 3.67 (s, 2H, CH_2), 2.35 (s, 6H, CH_3), 1.93 (h, *J* = 7.4 Hz, 2H, CH_2), 1.00 (t, *J* = 7.4 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 140.8, 139.9, 127.1, 125.5, 122.7, 122.7, 121.0, 120.3, 118.7, 108.7, 108.4, 64.5, 45.0 (2C), 44.7, 22.3, 11.8 ppm.

HRMS-ESI (m/z) found 222.1266 $[M]^+$, calcd 222.1277 for $[C_{16}H_{16}N]^+$ (loss of NMe_2).

Synthesis of 1-(9-Isopropyl-9*H*-carbazol-3-yl)-*N, N*-dimethylmethanamine (**2.26**)



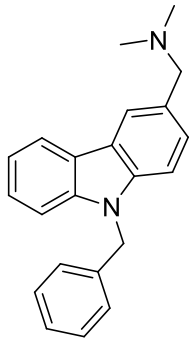
Synthesised according to general procedure A. 9-Isoropyl-*9H*-carbazole-3-carbaldehyde (498 mg, 2.1 mmol), dimethylamine hydrochloride (342 mg, 4.2 mmol), triethylamine (318 mg, 439 µL, 3.15 mmol), titanium isopropoxide (1.193 g, 1.243 mL, 4.2 mmol) and sodium borohydride (158 mg, 4.2 mmol). Product isolated as an orange oil, yield 29% (162 mg, 0.6 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.12 (d, *J* = 7.8 Hz, 1H, ar CH), 8.10 (s, 1H, ar CH), 7.54 – 7.49 (m, 5H, ar CH), 7.49 – 7.41 (m, 2H, ar CH), 7.25 – 7.18 (m, 2H, ar CH), 4.99 (p, *J* = 7.0 Hz, 1H, CH), 3.85 (s, 2H, CH_2), 2.46 (s, 6H, CH_3), 1.71 (d, *J* = 7.0 Hz, 6H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 139.8, 139.2, 127.1, 125.5, 123.3, 123.0, 121.6, 120.4, 118.7, 110.0, 63.7, 46.8, 44.1 (2C), 20.8 (2C) ppm.

HRMS-ESI (m/z) found 222.1265 $[M]^+$, calcd 222.127 for $[C_{16}H_{16}N]^+$ (loss of NMe_2).

Synthesis of 1-(9-Benzyl-9*H*-carbazol-3-yl)-*N, N*-dimethylmethanamine (**2.27**)



Synthesised according to general procedure A. 9-Benzyl-*9H*-carbazole-3-carbaldehyde (1.426 g, 5.0 mmol), dimethylamine hydrochloride (815 mg, 10.0 mmol), triethylamine (758 mg, 1.045 mL, 7.5 mmol), titanium isopropoxide (2.842 g, 2.960 mL, 10.0 mmol) and sodium borohydride (378 mg, 10.0 mmol). Product isolated as a yellow solid, yield 54% (850 mg, 2.7 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 2.32 (6H, s), 3.63 (2H, s), 5.52 (2H, s), 7.17-7.44 (12H, m), 8.08 (1H, s), 8.13 (1H, d, J= 8 Hz) ppm.

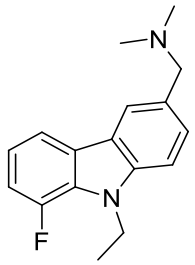
^{13}C NMR (126 MHz, Chloroform-*d*) δ 46.3, 46.6, 64.7, 108.6, 108.9, 119.1, 120.5, 121.0, 122.9, 123.0, 125.8, 126.4, 127.3, 127.4, 128.8, 137.2, 140.1, 141.1 ppm.

HRMS-ESI (m/z) found 270.1264 $[M]^+$, calcd 270.1277 for $[C_{20}H_{16}N]^+$ (loss of NMe_2).

General Procedure B

Representative procedure for reductive aminations using sodium triacetoxyborohydride

Synthesis of 1-(9-Ethyl-8-fluoro-9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**2.34**)



To a solution of 9-fluoro-9*H*-carbazole-3-carbaldehyde (**2.32**) (21 mg, 0.09 mmol) and methylamine (67 µL, 2M in THF, 0.13 mmol) in anhydrous THF (2 mL) was added sodium triacetoxyborohydride (46 mg, 0.21 mmol). The resulting suspension was stirred at room temperature for 18 h. THF was removed under a reduced pressure and to the residue was added DCM (10 mL) and H_2O (10 mL). The solution was filtered through a hydrophobic frit, solvent was removed in *vacuo* and crude product was purified by chromatography on silica gel (crude product loaded onto silica) using a gradient of 100% DCM to 9:1 DCM:MeOH over 15 minutes as an eluent. Product was isolated as a pale yellow solid, yield 89% (21 mg, 0.08 mmol).

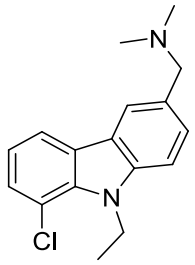
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.08 (s, 1H, ar CH), 7.86 (d, *J* = 7.4 Hz, 1H, ar CH), 7.55 (d, *J* = 8.4 Hz, 1H, ar CH), 7.43 (d, *J* = 8.4 Hz, 1H, ar CH), 7.21 – 7.09 (m, 2H, ar CH), 4.54 (q, *J* = 7.1 Hz, 2H, CH_2), 4.03 (s, 2H, CH_2), 2.54 (s, 6H, CH_3), 1.47 (t, *J* = 7.1 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 149.4 (d, $^1J_{FC}$ = 242.8 Hz), 140.5, 128.4, 128.0, 128.0, 126.3 (d, $^4J_{FC}$ = 4.8 Hz), 122.6, 119.3 (d, $^3J_{FC}$ = 6.4 Hz), 116.1 (d, $^4J_{FC}$ = 3.3 Hz), 112.0 (d, $^2J_{FC}$ = 18.5 Hz), 109.9, 109.0, 62.4, 42.8, 40.0 (d, $^4J_{FC}$ = 6.1 Hz), 15.0 ppm. Missing one quaternary carbon.

HRMS-ESI (m/z) found 226.1018 $[M]^+$, calcd 226.1027 for $[C_{15}H_{13}FN]^+$ (loss of NMe_2).

LC-MS purity 92% (UV), ret. time = 11.83 min.

Synthesis of 1-(9-Ethyl-8-chloro-9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**2.35**)



Synthesised according to general procedure B. 8-Chloro-9-ethyl-9*H*-carbazole-3-carbaldehyde (**2.33**) (11 mg, 0.045 mmol), methylamine (33 µL, 2M in THF, 0.067), anhydrous THF (2 mL), sodium triacetoxyborohydride (22 mg, 0.010 mmol). Product was isolated as an orange solid, yield 84% (11 mg, 0.038 mmol).

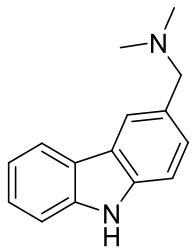
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.12 (s, 1H, ar CH), 8.00 (d, *J* = 7.7 Hz, 1H, ar CH), 7.59 (d, *J* = 8.4 Hz, 1H, ar CH), 7.45 (d, *J* = 8.7 Hz, 1H, ar CH), 7.42 (d, *J* = 8.0 Hz, 1H, ar CH), 7.14 (t, *J* = 7.7 Hz, 1H, ar CH), 4.79 (q, *J* = 7.2 Hz, 2H, CH_2), 4.01 (s, 2H, CH_2), 2.56 (s, 6H, CH_3), 1.49 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 140.9, 135.8, 128.4, 127.9, 125.7, 122.8, 122.2, 119.9, 118.9, 115.9, 109.2, 62.9, 43.3 (2C), 39.4, 15.6 ppm. Missing one quaternary carbon.

HRMS-ESI (m/z) found 242.0716 $[M]^+$, calcd 242.0731 for $[C_{15}H_{13}ClN]^+$ (loss of NMe_2).

LC-MS purity > 99% (UV), ret. time = 12.31 min.

Synthesis of 1-(9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**2.37**)



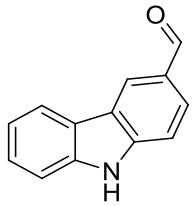
Synthesised according to general procedure B. 9*H*-carbazol-3-carbaldehyde (45 mg, 0.26 mmol), dimethylamine (130 µL, 2.0 M solution in THF, 0.26 mmol), anhydrous THF (5 mL), sodium triacetoxyborohydride (80 mg, 0.39 mmol). Product was isolated as a pale yellow solid, yield 88% (51 mg, 0.22 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.59 (br s, 1H, NH), 8.09 – 8.03 (m, 1H, ar CH), 8.02 (s, 1H, ar CH), 7.46 – 7.38 (m, 3H, ar CH), 7.38 – 7.33 (m, 1H, ar CH), 7.25 – 7.18 (m, 1H, ar CH), 3.77 (s, 2H, CH_2), 2.40 (s, 6H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 139.9, 139.2, 127.5, 126.8, 125.9, 123.3, 123.0, 121.4, 120.3, 119.4, 110.7, 110.6, 63.9, 44.2 (2C) ppm.

HRMS-ESI (*m/z*) found 180.0805 $[M]^+$, calcd 180.0808 for $[C_{13}H_{10}N]^+$ (loss of NMe_2).

Synthesis of 9*H*-carbazole-3-carbaldehyde (**2.36**)



Synthesised according to a known procedure.[530]

^{1}H NMR (500 MHz, DMSO-d_6) δ 11.82 (s, 1H, NH), 10.04 (s, 1H, CHO), 8.73 (s, 1H, ar CH), 8.32 – 8.16 (m, 1H, ar CH), 7.93 (d, J = 8.4 Hz, 1H, ar CH), 7.63 (d, J = 8.4 Hz, 1H, ar CH), 7.59 – 7.52 (m, 1H, ar CH), 7.52 – 7.43 (m, 1H, ar CH), 7.31 – 7.19 (m, 1H, ar CH) ppm.

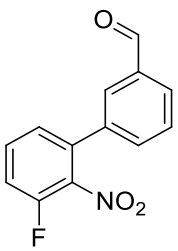
^{13}C NMR (126 MHz, DMSO-d_6) δ 191.9 (CHO), 149.4, 143.6, 140.5, 128.3, 126.7, 126.4, 122.6, 120.8, 119.9, 111.7, 111.5, 109.6 ppm.

HRMS-ESI (m/z) found 196.0756 $[M+H]^+$, calcd. 196.0757 for $[C_{13}H_9NO+H]^+$. The above data were identical to literature data.[531]

General Procedure C

Representative procedure for the Suzuki-Miyaura cross-coupling of 1-bromo-2-nitro-3-halogenated benzenes with 3-formylphenyl boronic acid

Synthesis of 3'-Fluoro-2'-nitro-[1,1'-biphenylbiphenyl]-3-carbaldehyde (**2.28)**



To a two-necked round bottomed flask was added 1-fluoro-3-iodo-2-nitrobenzene (2.000 g, 9.0 mmol), (3-formylphenyl)boronic acid (1.635 g, 10.9 mmol) and

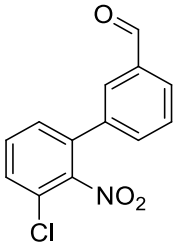
$Pd(dppf)Cl_2.CH_2Cl_2$ (371 mg, 0.45 mmol) . The flask was evaculated and backfilled with argon 3 times and anhydrous toluene (10 mL) was added. A solution of potassium carbonate (2.512 g, 18.1 mmol) in degassed H_2O (10 mL) was added and the resulting solution was stirred at 90 °C until TLC indicated completion (4/1 hexane/EtOAc). The solution was cooled and filtered through celite. EtOAc (50 mL) and H_2O (50 mL) were added and crude product was extracted with EtOAc (3 X 30 mL). The organic extracts were dried over anhydrous $MgSO_4$, filtered and the filtrate was concentrated under a reduced pressure to afford crude product that was purified by chromatography on silica gel (EtOAc/hexane 1/9 to 3/7 over 25 minutes). Product was isolated as a pale yellow crystalline solid, yield 64% (1.418 g, 5.7 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.06 (s, 1H, CHO), 7.96 (t, *J* = 4.4 Hz, 1H, ar CH), 7.89 (s, 1H, ar CH), 7.65 – 7.62 (m, 2H, ar CH), 7.59 (td, *J* = 8.1, 5.4 Hz, 1H, ar CH), 7.33 (t, *J* = 8.8 Hz, 1H, ar CH), 7.28 (d, *J* = 8.8 Hz, 1H, ar CH) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.3, 153.8 (d, $^1J_{FC}$ = 258.4 Hz), 139.1 (d, $^2J_{FC}$ = 19.2 Hz), 136.9, 136.2 (d, $^4J_{FC}$ = 2.0 Hz), 135.4, 133.5, 131.9 (d, $^3J_{FC}$ = 8.3 Hz), 130.1, 129.7, 129.1, 126.3 (d, $^4J_{FC}$ = 3.5 Hz), 166.6 (d, $^3J_{FC}$ =19.1 Hz) ppm.

MS-EI (m/z) found 245 $[M]^+$, calcd 245.

Synthesis of 3'-Chloro-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**2.29**)



Synthesised according to general procedure C. 1-Bromo-3-chloro-2-nitrobenzene (500 mg, 2.11 mmol), (3-formylphenyl)boronic acid (380 mg, 2.53 mmol), $Pd(dppf)Cl_2.CH_2Cl_2$ (86 mg, 0.10 mmol), potassium carbonate (584 mg, 4.22 mmol), anhydrous toluene (2.5 mL), degassed H_2O (2.5 mL). Product was isolated as white crystalline solid, yield 92 % (514 mg, 1.96 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.06 (s, 1H, CHO), 8.00 – 7.94 (m, 1H, ar CH), 7.88 (s, 1H ar CH), 7.66 – 7.61 (m, 2H, ar CH), 7.61 – 7.50 (m, 2H, ar CH), 7.41 – 7.35 (m, 1H, ar CH) ppm.

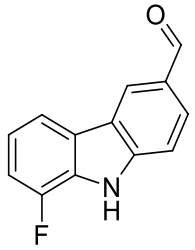
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.3, 148.0, 136.8, 136.3, 134.7, 133.6, 130.9, 130.1, 130.1, 129.6, 129.3, 129.3, 125.6 ppm.

MS-EI (m/z) found 261 $[M]^+$, calcd 261.

General Procedure D

Representative procedure for the reductive cyclization of intermediates **2.28** and **2.29**

Symthesis of 9-Fluoro-9*H*-carbazole-3-carbaldehyde (**2.30**)



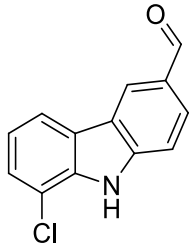
To a 15 mL microwave vial was added 3'-Fluoro-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**2.28**) (1.000 g, 4.07 mmol), triphenylphosphine (2.674 g, 10.19 mmol) and DMA (2 mL). The resulting solution was stirred under microwave irradiation at 200 °C for 2 h. The solution was cooled and diluted with EtOAc (50 mL) and H_2O (50 mL). Crude product was extracted with EtOAc (3 X 30 mL). The organic extracts were washed with brine (3 X 50 mL) and dried over anhydrous $MgSO_4$. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield crude product that was purified by column chromatography on silica gel (DCM/hexane gradient from 3/7 to 6/4 over 30 minutes). Product was isolated as an orange solid, yield 34% (298 mg, 1.39 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 12.28 (s, 1H, NH), 10.03 (s, 1H, CHO), 8.75 (s, 1H, ar CH), 8.07 (d, J = 7.8 Hz, 1H, ar CH), 7.96 (d, J = 8.4 Hz, 1H), 7.63 (d, J = 8.4 Hz, 1H, ar CH), 7.36 – 7.28 (m, 1H, ar CH), 7.25 – 7.18 (m, 1H, ar CH) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.3, 149.1 (d, $^1J_{FC}$ = 243.1 Hz), 144.2, 129.1, 128.50 (d, $^2J_{FC}$ = 13.8 Hz), 127.4, 126.69 (d, $^3J_{FC}$ = 5.2 Hz), 125.30 (d, $^2J_{FC}$ = 12.3 Hz), 122.9 (d, $^4J_{FC}$ = 2.6 Hz), 117.32 (d, $^3J_{FC}$ = 7.2 Hz), 112.38 (d, $^4J_{FC}$ = 4.4 Hz), 122.2, 112.1 ppm.

MS-EI (m/z) found 213 $[M]^+$, calcd 213.

Synthesis of 8-Chloro-9*H*-carbazole3-carbaldehyde (**2.31**)

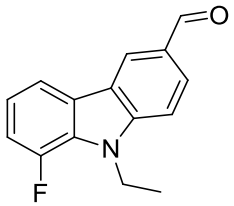


Synthesised according to general procedure D. 3'-Chloro-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**2.29**) (769 mg, 2.9 mmol), triphenylphosphine (1.927 g, 7.3 mmol) and DMA (1 mL). Product was isolated as a pale orange solid, yield, 9% (62 mg, 0.26 mmol). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 12.12 (s, 1H, NH), 10.06 (s, 1H, CHO), 8.78 (s, 1H, ar CH), 8.25 (d, J = 7.7 Hz, 1H, ar CH), 7.99 (d, J = 8.6 Hz, 1H, ar CH), 7.69 (d, J = 8.4 Hz, 1H, ar CH), 7.56 (d, J = 7.7 Hz, 1H, ar CH), 7.28 (t, J = 7.8 Hz, 1H, ar CH) ppm. **MS-EI** (m/z) found 229 $[M]^+$, calcd 229.

General Procedure E

Representative procedure for the N-alkylation of intermediates **2.30** and **2.31**

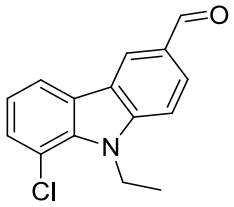
Synthesis of 9-Fluoro-9*H*-carbazole-3-carbaldehyde (**2.32**)



To a solution of 9-Fluoro-9*H*-carbazole-3-carbaldehyde (**2.30**) (160 mg, 0.78 mmol) in anhydrous DMF (3 mL) was added sodium hydride (62 mg, 60 % suspension in mineral oil, 1.56 mmol). The resulting suspension was stirred at room temperature for 30 minutes. Ethyl iodide (120 μL, 1.56 mmol) was added and the solution was stirred for 18 hours. The solution was diluted by EtOAc (20 mL) and H_2O (20 mL) and crude product was extracted with EtOAc (3 X 20 mL). The organic extracts were washed with brine (3 X 20 mL), dried over anhydrous $MgSO_3$ and filtered. The filtrate was concentrated under a reduced pressure and the resulting crude product was purified by chromatography on silica gel (DCM/hexane 2/8 to 1/1 over 20 minutes) to yield product as a pale yellow solid, yield 44% (80 mg, 0.33 mmol) that was carried on to the next step.

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.10 (s, 1H, CHO), 8.58 (s, 1H, ar CH), 8.07 – 8.01 (m, 1H, ar CH), 7.94 – 7.86 (m, 1H, ar CH), 7.52 – 7.45 (m, 1H, ar CH), 7.25 – 7.16 (m, 2H, ar CH), 4.56 (q, J = 7.2 Hz, 2H, CH_2), 1.50 (t, J = 7.2 Hz, 3H, CH_3) ppm.

Synthesis of 8-Chloro-9-ethyl-9*H*-carbazole-3-carbaldehyde (**2.33**)

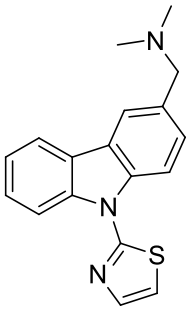


Synthesised according to general procedure E. 8-Chloro-9*H*-carbazole3-carbaldehyde (**2.31**) (60 mg, 0.26 mmol), anhydrous DMF (1 mL), sodium hydride (20 mg, 60 % suspension in mineral oil, 0.52 mmol), ethyl iodide (42 µL, 0.52 mmol). Product was isolated as a beige solid, yield 70% (47 mg, 0.016 mmol) that was carried on to the next step.

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.12 (s, 1H, CHO), 8.60 (s, 1H, ar CH), 8.09 – 8.02 (m, 2H, ar CH), 7.56 – 7.52 (m, 1H, ar CH), 7.49 (d, *J* = 7.7 Hz, 1H, ar CH), 7.23 (t, *J* = 7.7 Hz, 1H, ar CH), 4.85 (q, *J* = 7.1 Hz, 2H, CH_2), 1.53 (t, *J* = 7.1 Hz, 3H, CH_3) ppm.

^{13}C NMR (100 MHz, Chloroform-*d*) δ 191.4, 129.2, 128.6, 127.7, 126.1, 125.9, 123.7, 122.7, 120.9, 119.1, 116.3, 109.9, 109.1, 39.8, 15.6 ppm.

Synthesis of *N,N*-dimethyl-1-(9-(thiazol-2-yl)-9*H*-carbazol-3-yl)methanamine (**2.38**)



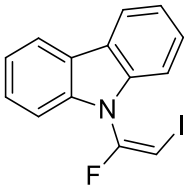
To a microwave vial containing 1-(9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**2.37**) (100 mg, 0.44 mmol) and Cs_2CO_3 (145 mg, 0.44 mmol) under argon was added anhydrous DMF (2 mL), 2-bromothiazole (43 µL, 0.48 mmol) and CuI (8 mg, 0.04 mmol). The vial was purged with argon for 5 minutes before heating under microwave irradiation at 220 °C for 40 minutes. The solution was cooled and diluted with saturated ammonium chloride (30 mL) and EtOAc (30 mL). Crude product was extracted with EtOAc (2 X 30 mL), the organic extracts were washed with brine (3 X 30 mL) and dried over anhydrous $MgSO_4$. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield crude product that was purified by chromatography on silica gel using

DCM/MeOH 1/0 to 9/1 over 20 minutes as an eluent to yield product as a brown oil, yield 53% (72 mg, 0.23 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.07 – 8.03 (m, 2H, ar CH), 7.93 (s, 1H, ar CH), 7.43 – 7.40 (m, 3H, ar CH), 7.38 – 7.35 (m, 1H, ar CH), 7.32 – 7.28 (m, 1H, ar CH), 7.01 – 6.98 (m, 1H, ar CH), 4.16 (s, 2H, CH_2), 3.04 (s, 6H, CH_3) ppm.
MS-ESI (m/z) found 308.050 (LC-MS) $[M+H]^+$, calcd 308.118 for $[C_{18}H_{17}N_3S+H]^+$.
LC-MS purity = 89 % (UV), ret time = 12.60 min.

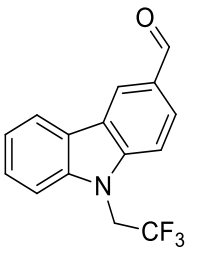
6.2 Chapter 3

Synthesis of (E)-9-(1-fluoro-2-iodovinyl)-9*H*-carbazole (**3.8**)



Synthesised according to general procedure E using modified conditions withcarbazole (84 mg, 0.5 mmol), 1,1,1-trifluoro-2-iodoethane (99 μL, 1 mmol), sodium hydride (40 mg, 60% suspension in mineral oil, 1 mmol), anhydrous DMF (1 mL). The solution was heated at 150 °C for 30 minutes under microwave irradiation. Purified by trituration with hexane to yield product as a white crystalline solid, yield 69% (116 mg, 0.34 mmol). The structure was confirmed by x-ray crystallography.
^{1}H NMR (600 MHz, Chloroform-*d*) δ 8.07 (d, *J* = 7.8 Hz, 2H, ar CH), 7.50 (ddd, *J* = 8.2, 7.2, 1.2 Hz, 2H, ar CH), 7.47 – 7.42 (m, 2H, ar CH), 7.38 – 7.33 (m, 2H, ar CH), 6.49 (d, *J* = 4.5 Hz, 1H, CH) ppm.
MS-EI found 337 $[M]^+$, calcd 337.

Synthesis of 9-(2,2,2-trifluoroethyl)-9*H*-carbazole-3-carbaldehyde (**3.9a**)



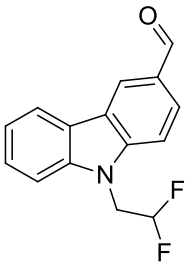
Synthesised according to general procedure E using modified conditions. Reaction was conducted under microwave irradiation at 150 °C for 30 minutes using 2,2,2-trifluoroethyl p-toluenesulfonate as an electrophile and **2.36** as a starting material. Reaction carried out on a 0.48 mmol scale. Pale yellow solid, yield 50% (68 mg, 0.24 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.07 (s, 1H, CHO), 8.75 (s, 1H, ar CH), 8.28 (d, J = 8.3 Hz, 1H, ar CH), 8.02 (d, J = 8.5 Hz, 1H, ar CH), 7.90 (d, J = 8.5 Hz, 1H, ar CH), 7.80 (d, J = 8.3 Hz, 1H, ar CH), 7.61 – 7.49 (m, 1H, ar CH), 7.40 – 7.25 (m, 1H, ar CH), 5.51 (q, $^{3}J_{FH}$ = 9.3 Hz, 2H, CH_2) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.3, 144.2, 141.3, 129.8, 127.5, 127.4, 125.4 (q, $^{1}J_{FC}$ = 281.8 Hz), 124.0 123.3, 122.9, 121.5, 121.2, 110.9, 110.8, 44.1 (q, $^{2}J_{FC}$ = 33.5 Hz) ppm.

HRMS-ESI (m/z) found 278.0789 $[M+H]^+$, calcd 278.0787 for $[C_{15}H_{10}F_3NO+H]^+$.

Synthesis of 9-(2,2-difluoroethyl)-9*H*-carbazole-3-carbaldehyde (**3.9b**)



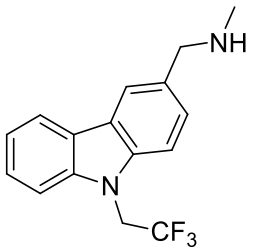
Synthesised according to general procedure E using modified conditions. Reaction was conducted under microwave irradiation at 150 °C for 30 minutes using 1-iodo-2,2-difluoroethane as an electrophile using **2.36** as a starting material. Reaction carried out on a 1.3 mmol scale. Yellow crystalline solid, yield 44% (115 mg, 0.44 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.06 (s, 1H, CHO), 8.73 (s, 1H, ar CH), 8.33 – 8.17 (m, 1H, ar CH), 7.99 (d, J = 8.4 Hz, 1H, ar CH), 7.81 (d, J = 8.4 Hz, 1H, ar CH), 7.76 – 7.66 (m, 1H, ar CH), 7.60 – 7.47 (m, 1H, ar CH), 7.36 – 7.24 (m, 1H, ar CH), 6.50 (tt, $^{2}J_{FH}$ = 54.5 Hz, $^{3}J_{HH}$ = 3.2 Hz, 1H, CH), 4.97 (td, $^{3}J_{FH}$ = 16.1 Hz, $^{3}J_{HH}$ = 3.2 Hz, 2H, CH_2) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.3, 144.6, 141.7, 129.4, 127.3, 127.2, 124.0, 123.0, 122.8, 121.1, 121.0, 115.1 (t, $^{1}J_{FC}$ = 241.9 Hz), 110.9, 110.8, 45.0 (t, $^{2}J_{FC}$ = 24.5 Hz) ppm.

HRMS-ESI (*m/z*) found 260.0884 $[M+H]^+$, calcd 260.0881 for $[C_{15}H_{11}F_2NO+H]^+$.

Synthesis of 1-(9-(2,2,2-trifluoroethyl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**3.7b**)



Synthesised according to general procedure A using **3.81** as a starting material and methylamine hydrochloride as an amine. Reaction carried out on a 0.31 mmol scale. Product isolated as a pale yellow solid, yield 40% (37 mg, 0.12 mmol).

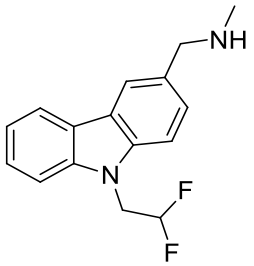
^{1}H NMR (500 MHz, Acetonitrile-d_3) δ 8.20 – 8.14 (m, 1H, ar CH), 8.12 (s, 1H, ar CH), 7.64 – 7.58 (m, 1H, ar CH), 7.58 – 7.47 (m, 3H, ar CH), 7.36 – 7.28 (m, 1H, ar CH), 5.07 (q, $^{3}J_{FH}$ = 9.1 Hz, 2H, CH_2), 3.89 (s, 2H, CH_2), 2.43 (s, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) 140.9, 140.2, 128.3, 127.7, 126.6, 126.1 (q, $^{1}J_{FC}$ = 281.2 Hz), 122.9 122.8, 121.3, 120.6, 120.5, 110.4, 110.1, 53.9, 44.05 (q, $^{2}J_{FC}$ = 33.4 Hz), 34.1 ppm.

LC-MS purity = 97% (UV), ret. time = 11.55 min.

HRMS-ESI (*m/z*) found 293.1259 $[M+H]^+$, calcd 293.1260 for $[C_{16}H_{15}F_3N_2+H]^+$.

Synthesis of 1-(9-(2,2-difluoroethyl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**3.7c**)



Synthesised according to general procedure A using **3.82** as a starting material and methylamine hydrochloride as an amine. Reaction carried out on a 0.28 mmol scale. Product isolated as a yellow oil, yield 58% (46 mg, 0.16 mmol).

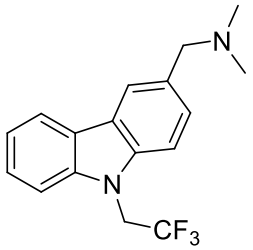
^{1}H NMR (500 MHz, Acetonitrile-d_3) δ 8.19 – 8.13 (m, 1H, ar CH), 8.10 (s, 1H, ar CH), 7.62 – 7.56 (m, 1H, ar CH), 7.56 – 7.45 (m, 3H, ar CH), 7.32 – 7.25 (m, 1H, ar CH), 6.31 (tt,$^{2}J_{FH}$ = 55.0 Hz, $^{3}J_{HH}$ = 3.2 Hz, 1H, CH), 4.76 (td, $^{3}J_{FH}$ = 15.6 Hz, $^{3}J_{HH}$ = 3.2 Hz, 2H, CH_2), 3.89 (s, 2H, CH_2), 2.43 (s, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Acetonitrile-d_3) δ 141.1, 139.9, 132.5, 126.6, 125.9, 122.9, 122.8, 120.1, 119.7, 119.6, 114.8 (t, $^{1}J_{FC}$ = 241.9 Hz), 109.3, 109.0, 55.7, 44.9 (t, $^{2}J_{FC}$ = 25.6 Hz), 35.1 ppm.

LC-MS purity = 96% (UV), ret. time = 11.01 min.

HRMS-ESI (*m/z*) found 275.1353 $[M+H]^+$, calcd 275.1354 for $[C_{16}H_{16}F_2N_2+H]^+$.

Synthesis of 1-(9-(2,2,2-trifluoroethyl)-9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**3.7a**)



Synthesised according to general procedure A using **3.81** as a starting material and methylamine hydrochloride as an amine. Reaction carried out on a 0.10 mmol scale. Product isolated as a white solid, yield 56% (17 mg, 0.05 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.04 – 8.00 (m, 1H, ar CH), 7.99 (s, 1H, ar CH), 7.45 – 7.39 (m, 2H, ar CH), 7.37 – 7.29 (m, 2H, ar CH), 7.27 – 7.20 (m, 1H, ar CH), 4.73 (q, $^{3}J_{FH}$ = 8.7 Hz, 2H, CH_2), 3.62 (s, 2H, CH_2), 2.28 (s, 6H, CH_3) ppm.

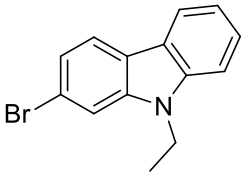
^{13}C NMR (126 MHz, Chloroform-*d*) δ 140.9, 140.1, 129.8, 127.8, 126.4, 124.3 (q, $^{1}J_{FC}$ = 281.5 Hz), 123.6, 123.4, 121.3, 120.6, 120.4, 108.7, 108.5, 45.3 (q, $^{2}J_{FC}$ = 35.7 Hz), 44.9 (2C) ppm.

LC-MS purity > 99% (UV), ret. time = 11.68 min.

HRMS-ESI (*m/z*) found 307.1419 $[M+H]^+$, calcd 307.1417 for $[C_{16}H_{17}F_3N_2+H]^+$.

6.3 Chapter 4

Synthesis of 2-bromo-9-ethyl-9*H*-carbazole (**4.5**)



Synthesised according to general procedure E. 2-bromocarbazole (5.000 g, 20.31 mmol), anhydrous DMF (20 mL), sodium hydride (1.625 g, 40.63 mmol), ethyl iodide (3.266 mL, 40.63 mmol). Crude product was isolated as a beige crystalline solid and assessed to

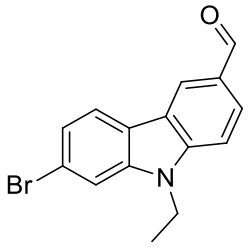
have a purity > 95% (^{1}H NMR), hence was not purified using the procedure outline in general procedure E. Yield, quantitative (5.53 g, 20.17 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.08 (d, *J* = 6.8 Hz, 1H, ar CH), 7.95 (d, *J* = 8.2 Hz, 1H, ar CH), 7.57 (s, 1H, ar CH), 7.51 (t, *J* = 8.2, 7.2, 1.1 Hz, 2H, ar CH), 7.42 (d, *J* = 8.2 Hz, 1H, ar CH), 7.35 (d, *J* = 8.2, 1.6 Hz, 1H, ar CH), 7.27 (t, 2H, ar CH), 4.33 (q, *J* = 7.2 Hz, 2H, CH_2), 1.45 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 140.7, 140.0, 126.0, 122.4, 121.9, 121.5, 120.4, 119.3, 119.2, 111.5, 108.6, 37.6, 13.7 ppm.

HRMS-ESI (m/z) found 274.0226 $[M+H]^+$, calcd 274.0226 for $[C_{14}H_{12}BrN+H]^+$.

Synthesis of 7-bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**)



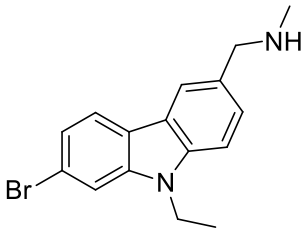
To a flask of anhydrous DMF (10 mL) under argon at 0 °C was added phosphorous (V) oxychloride (5.963 mL, 63.97 mmol) dropwise. The resulting solution was allowed to warm to room temperature. To the resulting solution was added a solution of 2-bromo-9-ethyl-9*H*-carbazole (**4.5**) (5.846 g, 21.32 mmol) in anhydrous DMF (10 mL) dropwise and the resulting solution was heated at 80 °C for 18 h. The solution was cooled to room temperature and was poured into ice water (50 mL). The solution was neutralized by addition of 10% w/v $NaOH_{(aq)}$. Crude product was extracted with EtOAc (3 X 100 mL). The organic extracts were washed with brine (3 X 50 mL), dried over anhydrous $MgSO_4$ and concentrated under a reduced pressure to yield crude product that was purified by column chromatography on silica gel (EtOAc/hexane 0/1 to 1/4). Product was isolated as a beige flocculent solid, yield 46% (3.00 g, 9.9 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.08 (s, 1H, CHO), 8.55 (s, 1H, ar CH), 8.02 (d, *J* = 8.5 Hz, 1H, ar CH), 7.97 (d, *J* = 8.3 Hz, 1H, ar CH), 7.59 (s, 1H, ar CH), 7.46 (d, *J* = 8.5 Hz, 1H, ar CH), 7.41 (d, *J* = 8.3 Hz, 1H, ar CH), 4.34 (q, *J* = 7.3 Hz, 2H, CH_2), 1.45 (t, *J* = 7.3 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.6, 143.6, 141.4, 128.9, 127.5, 123.8, 123.4, 122.6, 121.9, 121.8, 120.3, 112.3, 108.9, 38.0, 13.7 ppm.

HRMS-ESI (m/z) found 302.0180 $[M+H]^+$, calcd 302.0175 for $[C_{15}H_{12}BrNO+H]^+$.

Synthesis of 1-(7-bromo-9-ethyl-9*H*-carbzol-3-yl)-*N*-methylmethanamine (**4.3a**)



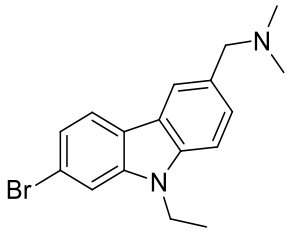
Synthesised according to general procedure A. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (50 mg, 0.17 mmol), methylamine hydrochloride (23 mg, 0.34 mmol), trimethylamine (36 µL, 0.26 mmol), titanium (IV) isopropoxide (101 µL, 0.34 mmol), sodium borohydride (13 mg, 0.34 mmol), DCM (1.5 mL), EtOH (1.5 mL). Product was isolated as a white solid, yield 52% (28 mg, 0.088 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.06 (s, 1H, ar CH), 7.91 (d, *J* = 8.2 Hz, 1H, ar CH), 7.52 – 7.47 (m, 2H, ar CH), 7.35 – 7.29 (m, 2H, ar CH), 4.26 (q, *J* = 7.2 Hz, 2H, CH_2), 3.98 (s, 2H, CH_2), 2.52 (s, 3H, CH_3), 1.38 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

LC-MS purity = 98 % (UV), ret. time = 6.81 min.

MS-ESI (m/z) found 286.1 $[M]^+$ (LC-MS), calcd 286.2 (loss of NHMe).

Synthesis of 1-(9-bromo-9-ethyl-9*H*-carbazol-3-yl)-*N,N*-dimethylmethanamine (**4.3b**)



Synthesised according to general procedure B. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (700 mg, 2.32 mmol), dimethylamine (1.165 mL, 2 M solution in THF, 2.32 mmol), anhydrous THF (15 mL), sodium triacetoxyborohydride (738 mg, 3.48 mmol). Product was isolated as a pale orange oil, yield 84% (647 mg, 1.95 mmol).

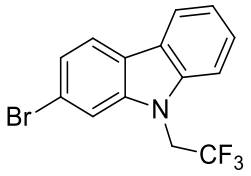
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.02 (s, 1H, ar CH), 7.93 (d, *J* = 8.2 Hz, 1H, ar CH), 7.55 (s, 1H, ar CH), 7.48 (d, *J* = 8.2 Hz, 1H, ar CH), 7.36 (d, *J* = 8.4 Hz, 1H, ar CH), 7.33 (d, *J* = 8.2 Hz, 1H, ar CH), 4.30 (q, *J* = 7.1 Hz, 2H, CH_2), 3.69 (s, 2H, CH_2), 2.36 (s, 6H, CH_3), 1.43 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 141.0, 139.5, 128.7, 127.6, 122.4, 121.9, 121.7, 121.6, 121.1, 119.2, 111.5, 108.4, 64.3, 44.9 (2C), 37.7, 13.7 ppm.

HRMS-ESI (m/z) found 286.0216 $[M]^+$, calcd 286.0226 for $[C_{15}H_{13}BrN]^+$ (loss of NMe_2).

LC-MS purity = 95% (UV), ret. time = 12.50 min.

Synthesis of 2-bromo-9-(2,2,2-trifluoroethyl)-9*H*-carbazole (**4.6**)



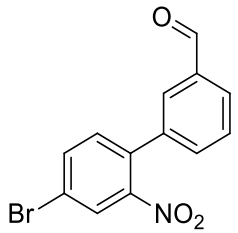
Synthesised according to general procedure E using modified conditions. Reaction was conducted under microwave irradiation at 150 °C for 30 minutes using 2,2,2-trifluoroethyl p-toluenesulfonate as an electrophile. 2-Bromocarbazole (719 mg, 2.92 mmol), 2,2,2-trifluoroethyl p-toluenesulfonate (1.4263 g, 5.84 mmol), cesium carbonate (1.9040 g, 5.84 mmol), anhydrous DMF (10 mL). Purification by chromatography on silica gel using DCM/ hexane 0/1 to 1/9 over 20 minutes as an eluent. Product was isolated as a white solid, yield 36% (347 mg, 1.05 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.05 (d, *J* = 8.1 Hz, 1H, ar CH), 7.92 (d, *J* = 8.2 Hz, 1H, ar CH), 7.56 (s, 1H, ar CH), 7.52 (t, *J* = 7.7 Hz, 1H, ar CH), 7.43 – 7.39 (m, 2H, ar CH), 7.32 (t, *J* = 7.5 Hz, 1H, ar CH), 4.74 (q, $^3J_{FH}$ = 8.6 Hz, 2H, CH_2) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 141.3, 140.5, 126.75, 123.6, 123.51 (q, $^1J_{FC}$ = 280.7 Hz), 122.9, 122.5, 121.5, 120.8, 120.4, 119.8, 111.9, 108.8, 45.2 (q, $^2J_{FC}$ = 35.6 Hz) ppm.

MS-EI (m/z) found 327 $[M]^+$, cacld 327.

Synthesis of 4'-bromo-2'-nitro-[1,1'-biphenyl]-3-carbalehyde (**4.9**)



Synthesised according to general procedure C. 4-Bromo-1-iodo-2-nitrobenzene (4.000 g, 12.19 mmol), (3-formylphenyl)boronic acid (2.012 g, 13.41 mmol), potassium carbonate (3.372 g, 24.39 mmol), $Pd(dppf)Cl_2.CH_2Cl_2$ (498 mg, 0.60 mmol), anhydrous toluene (20

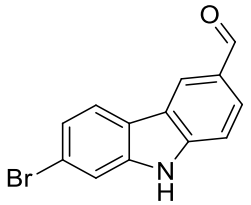
mL), degassed H_2O (10 mL). Crude product was purified by chromatography on silica gel using DCM/hexane 0/1 to 1/1 over 30 minutes as an eluent. Product was isolated as an orange oil, yield 84% (3.1592 g, 10.32 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.05 (s, 1H, CHO), 8.11 (s, 1H, ar CH), 7.94 (d, *J* = 8.4 Hz, 1H, ar CH), 7.84 – 7.78 (m, 2H, ar CH), 7.61 (t, *J* = 7.6 Hz, 2H, ar CH), 7.54 (d, *J* = 8.5 Hz, 1H, ar CH), 7.34 (d, *J* = 8.2 Hz, 1H, ar CH) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.3, 149.1, 137.6, 136.8, 135.7, 134.0, 133.6, 133.1, 129.8, 129.4, 128.7, 127.4, 122.2 ppm.

HRMS-ESI (m/z) found 305.9757 $[M+H]^+$, calcd. 305.9760 for $[C_{13}H_8BrNO_3+H]^+$.

Synthesis of 7-bromo-9*H*-carbazole-3-carbaldehyde (**4.8**)



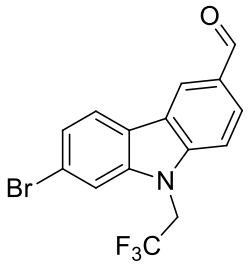
Synthesised according to general procedure D using modified conditions. Reaction was stirred at 165 °C for 18 h. 4'-Bromo-2'-nitro-[1,1'-biphenyl]-3-carbalehyde (**4.9**) (3.1592 g, 10.32 mmol), triphenylphosphine (6.7671 g, 25.8 mmol), DMA (10 mL). Crude product was purified by chromatography on silica gel using EtOAc/hexane 0/1 to 1/4 over 40 minutes. Product was isolated as an orange solid, yield 22% (629 mg, 2.29 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 11.94 (s, 1H, NH), 10.02 (s, 1H, CHO), 8.72 (s, 1H, ar CH), 8.18 (d, *J* = 8.3 Hz, 1H, ar CH), 7.94 (d, *J* = 8.4 Hz, 1H, ar CH), 7.74 (s, 1H, ar CH), 7.63 (d, *J* = 8.4 Hz, 1H, ar CH), 7.37 (d, *J* = 8.3 Hz, 1H, ar CH) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.3, 144.1, 141.8, 129.1, 127.2, 124.8, 123.1, 122.9, 122.4, 122.1, 119.6, 114.7, 112.2 ppm.

MS-EI (m/z) found 273 $[M]^+$, calcd 273.

Synthesis of 7-bromo-9-(2,2,2-trifluoroethyl)-9*H*-carbazole-3-carbaldehyde (**4.7**)



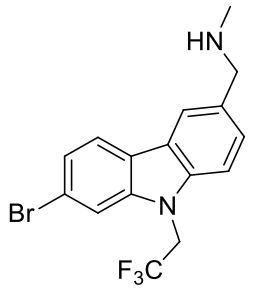
Synthesised according to general procedure E using modified conditions. Reaction was conducted under microwave irradiation at 150 °C for 30 minutes using 2,2,2-trifluoroethyl p-toluenesulfonate as an electrophile. 7-Bromo-9*H*-carbazole-3-carbaldehyde (**4.8**) (200 mg, 0.72 mmol), 2,2,2-trifluoroethyl p-toluenesulfonate (356 mg, 1.45 mmol), cesium carbonate (475 mg, 1.45 mmol), anhydrous DMF (3 mL). Crude product was purified by chromatography on silica gel using EtOAc/hexane ¼ to 1/1 over 20 minutes. Product was isolated as a white solid, yield 52% (134 mg, 0.37 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.06 (s, 1H, CHO), 8.78 (s, 1H, ar CH), 8.25 (d, J = 8.3 Hz, 1H, ar CH), 8.14 (s, 1H, ar CH), 8.05 (d, J = 8.6 Hz, 1H, ar CH), 7.91 (d, J = 8.6 Hz, 1H, ar CH), 7.49 (d, J = 8.3 Hz, 1H, ar CH), 5.55 (q, $^{3}J_{FH}$ = 9.2 Hz, 2H, CH_2) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.3, 144.3, 142.3, 130.2, 127.9, 125.3 (q, $^{1}J_{FC}$ = 281.7 Hz), 124.5, 124.2, 123.0, 122.7, 122.1, 120.4, 113.9, 11.2, 44.1 (q, $^{2}J_{FC}$ = 33.0 Hz) ppm.

MS-EI (m/z) found 355 $[M]^+$, calcd 355.

Synthesis of 1-(7-bromo-9-(2,2,2-trifluoroethyl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.10**)



Synthesised according to general procedure A. 7-Bromo-9-(2,2,2-trifluoroethyl)-9*H*-carbazole-3-carbaldehyde (**4.7**) (85 mg, 0.24 mmol), methylamine hydrochloride (32 mg, 0.48 mmol), Et_3N (49 µL, 0.35 mmol), titanium (IV) isopropoxide (141 µL, 0.48 mmol), DCM (1 mL), EtOH (1 mL), sodium borohydride (18 mg, 0.48 mmol). Product was isolated as a beige solid, yield 36% (32 mg, 0.08 mmol).

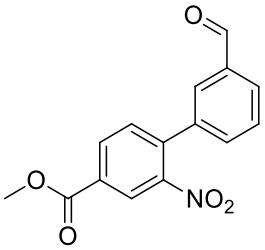
^{1}H NMR (500 MHz, DMSO-d_6) δ 8.32 (s, 1H, ar CH), 8.08 (s, 1H, ar CH), 8.05 (d, J = 8.3 Hz, 1H, ar CH), 7.80 (d, J = 8.4 Hz, 1H, ar CH), 7.69 (d, J = 8.4 Hz, 1H, ar CH), 7.44 (d, J = 8.3 Hz, 1H, ar CH), 5.49 (q, $^{3}J_{FH}$= 9.2 Hz, 2H, CH_2), 4.23 (s, 2H, CH_2), 2.53 (s, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 141.9, 140.9, 128.9, 124.5, 123.7, 123.4 (q, $^{1}J_{FC}$ = 285.3 Hz), 122.8, 122.4, 122.3, 121.8, 119.8, 113.6, 110.8, 51.9, 43.9 (q, $^{2}J_{FC}$ = 33.2 Hz), 32.2 ppm.

HRMS-ESI (m/z) found 339.9470 $[M]^+$, calcd 339.9943 for $[C_{15}H_{10}BrF_3N]^+$ (loss of NHMe).

LC-MS purity = 97% (UV), ret. Time = 12.55 min.

Synthesis of methyl-3'-formyl-2-nitro-[1,1'-biphenyl]-4-carboxylate (**4.11**)



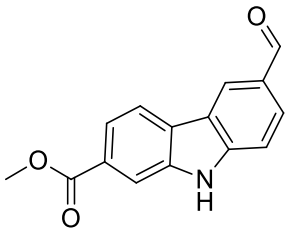
Synthesised according to general procedure C. Methyl-4-bromo-3-nitrobenzoate (260 mg, 1 mmol), (3-formylphenyl)boronic acid (165 mg, 1.1 mmol), potassium carbonate (276 mg, 2 mmol), Pd(dppf)Cl_2.CH_2Cl_2 (40 mg, 0.05 mmol), anhydrous toluene (2 mL), degassed H_2O (1 mL). Crude product was purified by chromatography on silica gel using DCM/hexane 0/1 to 1/1 as an eluent. Product was isolated as a white solid, yield 85% (242 mg, 0.85 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.07 (s, 1H, CHO), 8.59 (s, 1H, ar CH), 8.32 (d, *J* = 8.0 Hz, 1H, ar CH), 7.97 (d, *J* = 7.6 Hz, 1H, ar CH), 7.87 (s, 1H, ar CH), 7.64 (t, *J* = 7.6 Hz, 1H, ar CH), 7.57 (t, *J* = 8.3 Hz, 2H, ar CH), 4.02 (s, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.3, 164.6, 148.9, 139.0, 137.7, 136.8, 133.5, 133.1, 132.2, 131.2, 130.0, 129.4, 128.7, 125.5, 52.8 ppm.

MS-EI (m/z) found 285 $[M]^+$, calcd 285.

Synthesis of methyl-6-formyl-9*H*-carbazole-2-carboxylate (**4.12**)



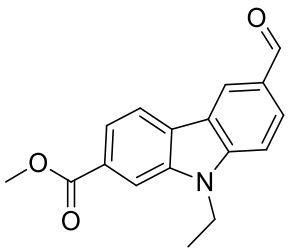
Synthesised according to general procedure D using modified conditions. Reaction was stirred at 165 °C for 18 h. Methyl-6-formyl-9*H*-carbazole-2-carboxylate (**4.12**) (1.300 g, 4.5 mmol), triphenyl phosphine (2.988 g, 11.4 mmol), DMA (3 mL). Crude product was purified by chromatography on silica gel using EtOAc/hexane 0/1 to 1/1 as an eluent. Product was isolated as a white solid, yield 34% (394 mg, 1.5 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 12.10 (s, 1H, NH), 10.04 (s, 1H, CHO), 8.81 (s, 1H, ar CH), 8.36 (d, J = 8.2 Hz, 1H, ar CH), 8.15 (s, 1H, ar CH), 7.99 (d, J = 8.5 Hz, 1H, ar CH), 7.85 (d, J = 8.2 Hz, 1H, ar CH), 7.69 (d, J = 8.5 Hz, 1H, ar CH), 3.89 (s, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.3, 167.0, 145.0, 140.2, 129.1, 127.7, 126.7, 125.9, 122.1, 121.2, 120.8, 117.7, 113.3, 112.4, 52.6 ppm.

MS-EI (m/z) found 253 $[M]^+$, calcd 253.

Synthesis of methyl-9-ethyl-6-formyl-9*H*-carbazole-2-carboxylate (**4.13**)



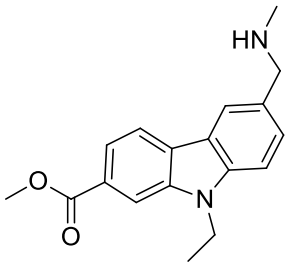
Synthesised according to general procedure E. Methyl-6-formyl-9*H*-carbazole-2-carboxylate (**4.12**) (78 mg, 0.30 mmol), anhydrous DMF (2 mL), sodium hydride (14 mg, 60% suspension in mineral oil, 0.34 mmol), ethyl iodide (27 µL, 0.34 mmol). Crude product was isolated as an orange solid and assessed to have a purity > 95% (^{1}H NMR). Yield, quantitative (86 mg, 0.30 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.08 (s, 1H, CHO), 8.85 (s, 1H, ar CH), 8.42 (d, J = 8.1 Hz, 1H, ar CH), 8.27 (s, 1H, ar CH), 8.07 (d, J = 8.6 Hz, 1H, ar CH), 7.91 (d, J = 8.1 Hz, 1H, ar CH), 7.86 (d, J = 8.6 Hz, 1H, ar CH), 4.60 (q, J = 7.1 Hz, 2H, CH_2), 3.93 (s, 3H, CH_3), 1.36 (t, J = 7.1 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.4, 167.4, 144.6, 140.1, 128.9, 128.1, 128.0, 126.7, 124.9, 122.3, 121.4, 120.4, 117.7, 110.9, 109.9, 109.1, 52.3, 38.1, 13.9 ppm.

MS-EI (m/z) found 281 $[M]^+$, calcd 281.

Synthesis of methyl-9-ethyl-6-((methylamino)methyl)-9*H*-carbazole-2-carboxylate (**4.14**)



Synthesised according to general procedure A. Methyl-9-ethyl-6-formyl-9*H*-carbazole-2-carboxylate (**4.13**) (50 mg, 0.17 mmol), methylamine hydrochloride (24 mg, 0.35 mmol), trimethylamine (35 µL, 0.25 mmol), titanium (IV) isopropoxide (104 µL, 0.35 mmol), DCM (2 mL), EtOH (2 mL), sodium borohydride (13 mg, 0.35 mmol). Product was isolated as a white solid, yield 16% (8 mg, 0.026 mmol).

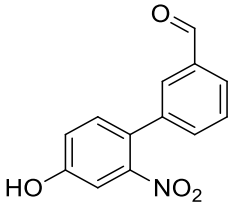
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.15 – 8.08 (m, 3H, ar CH), 7.92 (d, *J* = 8.1 Hz, 1H, ar CH), 7.52 (d, *J* = 8.4 Hz, 1H, ar CH), 7.40 (d, *J* = 8.4 Hz, 1H, ar CH), 4.41 (q, *J* = 7.2 Hz, 2H, CH_2), 3.99 (s, 2H, CH_2), 3.96 (s, 3H, CH_3), 2.53 (s, 3H, CH_3), 1.45 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 167.8, 140.7, 139.6, 127.8, 127.1, 126.5, 122.2, 121.1, 120.0, 120.0, 110.3, 109.9, 108.7, 55.7, 53.3, 52.0, 37.7, 13.8 ppm.

HRMS-ESI (m/z) found 266.1163 $[M]^+$, calcd 266.1176 for $[C_{17}H_{17}NO_2]^+$ (loss of NHMe).

LC-MS purity = 97% (UV), ret. time = 11.60 min.

Synthesis of 4'-hydroxy-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**4.15**)



To a solution of 4'-((*tert*-Butyldimethylsilyl)oxy)-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**4.18**) (2.515 g, 7.03 mmol) in THF (10 mL) at 0 °C was added *tert*-butylammonium fluoride (20.8 mL, 1M solution in THF, 20.68 mmol). The solution was allowed to warm to room temperature before it was diluted with H_2O (20 mL) and acidified with 1 N $HCl_{(aq)}$ to pH 4. Crude product was extracted with EtOAc (3 X 30

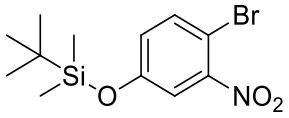
mL), washed with H_2O (30 mL) and dried over anhydrous $MgSO_4$. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield product as a yellow solid, yield 96% (1.657 g, 6.81 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.62 (br s, 1H, OH), 10.02 (s, 1H, CHO), 7.89 (d, J = 7.4 Hz, 1H, ar CH), 7.78 (s, 1H, ar CH), 7.63 (t, J = 7.5 Hz, 1H, ar CH), 7.59 (d, J = 7.7 Hz, 1H, ar CH), 7.39 (d, J = 8.4 Hz, 1H, ar CH), 7.36 (s, 1H, ar CH), 7.16 (d, J = 8.4 Hz, 1H, ar CH) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 193.3, 158.2, 149.4, 138.7, 136.8, 134.3, 133.4, 129.9, 129.2, 128.9, 125.1, 120.6, 111.2 ppm.

MS-EI (m/z) found 243 $[M]^+$, calcd 243.

Synthesis of (4-bromo-3-nitrophenoxy)(*tert*-butyl)dimethylsilane (**4.17**)



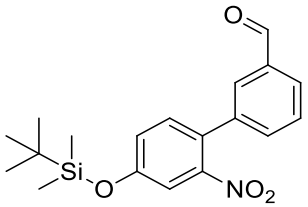
To a solution of 4-bromo-3-nitrophenol (436 mg, 2 mmol) and *tert*-butyldimethylsilyl chloride (331 mg, 2.2 mmol) in anhydrous DCM (4 mL) was added trimethylamine (418 µL, 3 mmol) at 0 °C. The solution was allowed to warm to room temperature and was stirred until TLC indicated completion (2 h). The suspension was poured into saturated sodium hydrogen carbonate and DCM was added. The solution was filtered through a hydrophobic frit and the organic extract was filtered through a silica plug, eluting product with DCM. Solvent was removed under a reduced pressure to yield product as a yellow oil, yield 98% (656 mg, 1.97 mmol).

^{1}H NMR (500 MHz, Chloroform-d) δ 7.55 (d, J = 8.7 Hz, 1H, ar CH), 7.30 (s, 1H, ar CH), 6.92 (d, J = 8.7 Hz, 1H, ar CH), 0.98 (s, 9H, CH_3), 0.24 (s, 6H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-d) δ 155.5, 135.4, 125.2, 117.2, 105.2, 25.4 (3C), 18.16, -4.53 (2C) ppm.

MS-EI (m/z) found 333 $[M]^+$, calcd 333.

Synthesis of 4'-((*tert*-butyldimethylsilyl)oxy)-2'-nitro-[1,1'biphenyl]-3-carbaldehyde (**4.18**)

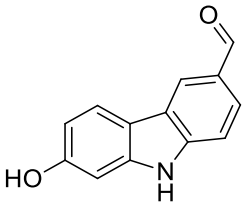


Synthesised according to general procedure C. (4-Bromo-3-nitrophenoxy)(*tert*-butyl)dimethylsilane (**4.17**) (384 mg, 1.15 mmol), (3-formylphenyl)boronic acid (190 mg, 1.27 mmol), potassium carbonate (317 mg, 2.3 mmol), $Pd(dppf)Cl_2.CH_2Cl_2$ (47 mg, 0.057 mmol), anhydrous toluene (2 mL) and H_2O (1 mL). Crude product was purified by chromatography on silica gel using DCM/hexane 0/1 to 1/1 as an eluent. Product was isolated as a yellow oil, yield 60% (248 mg, 0.69 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.04 (s, 1H, CHO), 7.89 (d, *J* = 7.2 Hz, 1H, ar CH), 7.81 (s, 1H, ar CH), 7.60 – 7.50 (m, 3H, ar CH), 7.41 (s, 1H, ar CH), 7.30 (d, *J* = 8.3 Hz, 1H, ar CH), 7.13 (d, *J* = 8.3 Hz, 1H, ar CH), 1.03 (s, 9H, CH_3), 0.29 (s, 6H, CH_3).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.6, 156.0, 149.3, 138.6, 136.7, 133.9, 132.7, 132.6, 129.2, 129.1, 129.0, 127.9, 124.3, 115.8, 25.5 (3C), 18.1, -4.4 (2C) ppm.

Synthesis of 7-hydroxy-9*H*-carbazole-3-carbaldehyde (**4.19**)



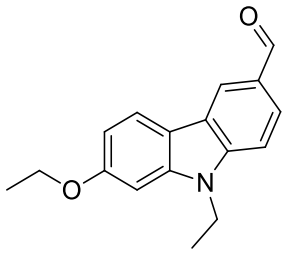
Synthesised according to general procedure D using modified conditions. Reaction was stirred at 165 °C for 18 h. 4'-Hydroxy-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**4.15**) (90 mg, 0.37 mmol), triphenylphosphine (242 mg, 0.92 mmol) and DMA (2 mL). Crude product was purified by chromatography on silica gel using EtOAc/hexane 0/1 to 1/1. Product was isolated as a white solid, yield 21% (17 mg, 0.08 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 11.53 (s, 1H, NH), 9.97 (s, 1H, CHO), 9.61 (s, 1H, OH), 8.50 (s, 1H, ar CH), 7.97 (d, *J* = 7.4 Hz, 1H, ar CH), 7.78 (d, *J* = 7.6 Hz, 1H, ar CH), 7.48 (d, *J* = 7.4 Hz, 1H, ar CH), 6.85 (s, 1H, ar CH), 6.70 (d, *J* = 7.6 Hz, 1H, ar CH) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.3, 157.7, 143.9, 142.6, 128.5, 125.6, 123.6, 121.9, 115.4, 111.1, 110.1, 97.3 ppm.

MS-EI (m/z) found 211 $[M]^+$, calcd 211.

Synthesis of 7-ethyoxy-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.20**)



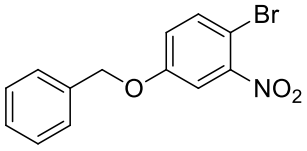
Synthesised according to general procedure E modifying the solvent according to Compain-Batissou *et al*.[532] 7-Hydroxy-9*H*-carbazole-3-carbaldehyde (**4.19**) (80 mg, 0.37 mmol), anhydrous THF (10 mL), anhydrous DMF (54 µL, 0.70 mmol), sodium hydride (36 mg, 60% dispersion in mineral oil, 0.92 mmol), ethyl iodide (32 µL, 0.40 mmol). Crude product was purified by chromatography on silica gel using EtOAc/hexane 1/4 to 1/1. Product was isolated as a pale yellow solid, yield 17% (17 mg, 0.063 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.03 (s, 1H, CHO), 8.61 (s, 1H, ar CH), 8.14 (d, J = 8.5 Hz, 1H, ar CH), 7.90 (d, J = 9.4 Hz, 1H, ar CH), 7.71 (d, J = 8.4 Hz, 1H, ar CH), 6.89 (d, J = 10.3 Hz, 1H, ar CH), 4.47 (q, J = 7.4 Hz, 2H, CH_2), 4.18 (q, J = 7.0 Hz, 2H, CH_2), 1.40 (t, J = 6.9 Hz, 3H, CH_3), 1.32 (t, J = 7.1 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.9, 129.9, 123.4, 122.7, 122.7, 121.5, 117.7, 108.8, 108.3, 105.6, 95.6, 94.4, 64.0, 37.8, 14.9, 13.6 ppm. Missing one quaternary carbon.

MS-EI (m/z) found 267 $[M]^+$, calcd 267.

Synthesis of 4-(benzyloxy)-1-bromo-2-nitrobenzene (**4.21**)



To a solution of 4-bromo-3-nitrophenol (800 mg, 3.66 mmol) in anhydrous THF (5 mL) was added sodium hydride (161 mg, 60% dispersion in mineral oil, 4.03 mmol). The resulting suspension was stirred for 30 minutes at room temperature. Benzyl bromide (478 µL, 4.03 mmol) was added and the resulting solution was stirred at 30 °C until TLC

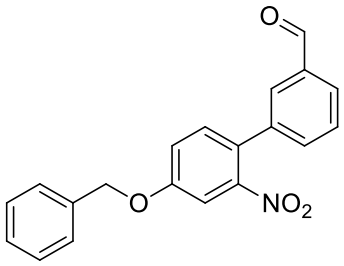
indicated completion (18 h). Solvent was removed under a reduced pressure and EtOAc (50 mL) and H_2O (50 mL) were added to the residue. Crude product was extracted with EtOAc (3 X 20 mL) and the organic extracts were washed with H_2O (30 mL) and dried over anhydrous $MgSO_4$. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield crude product that was purified by chromatograpy on silica gel using EtOAc/hexane 1/4 to 1/1 as an eluent. Product was isolated as a white solid, yield 80% (905 mg, 2.93 mmol).

^{1}H NMR (399 MHz, DMSO-d_6) δ 7.77 (d, J = 8.9 Hz, 1H, ar CH), 7.71 (d, J = 3.0 Hz, 1H, ar CH), 7.48 – 7.31 (m, 5H, ar CH), 7.26 (dd, J = 8.9, 3.0 Hz, 1H, ar CH), 5.18 (s, 2H, CH_2) ppm.

^{13}C NMR (126 MHz, Chloroform-d) δ 158.2, 150.1, 135.5, 135.3, 128.8 (2C), 128.5, 127.5 (2C), 120.5, 111.9, 104.8, 70.9 ppm.

MS-EI (m/z) found 307 $[M]^+$, calcd 307.

Synthesis of 4'-(benzyloxy)-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**4.22**)



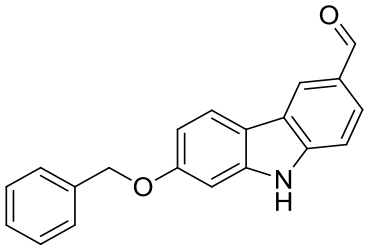
Synthesised according to general procedure C. 4-(Benzyloxy)-1-bromo-2-nitrobenzene (**4.21**) (616 mg, 2 mmol), (3-formylphenyl)boronic acid (329 mg, 2.2 mmol), potassium carbonate (552 mg, 4 mmol), $Pd(dppf)Cl_2.CH_2Cl_2$ (81 mg, 0.1 mmol), anhydrous toluene (2 mL) and degassed H_2O (1 mL). Crude product was purified by recrystallization in EtOAc/hexane. Product was isolated as a beige crystalline solid, yield 54% (528 mg, 1.64 mmol).

^{1}H NMR (500 MHz, Chloroform-d) δ 10.06 (s, 1H, CHO), 7.91 (d, J = 8.7 Hz, 1H, ar CH), 7.82 (s, 1H, ar CH), 7.62 – 7.51 (m, 3H, ar CH), 7.50 – 7.41 (m, 3H, ar CH), 7.42 – 7.33 (m, 1H, ar CH), 7.30 – 7.23 (m, 1H, ar CH), 5.19 (s, 2H, CH_2) ppm.

^{13}C NMR (126 MHz, Chloroform-d) δ 191.6, 158.7, 149.3, 138.6, 136.7, 135.5, 134.0, 132.8, 129.1, 129.1, 128.8 (2C), 128.4, 127.6, 127.5 (2C), 119.6, 110.5, 70.8 ppm. Missing one quarternary carbon.

HRMS-ESI (m/z) found 356.0892 $[M+Na]^+$, calcd 356.0893 for $[C_{20}H_{15}NO_4+Na]^+$.

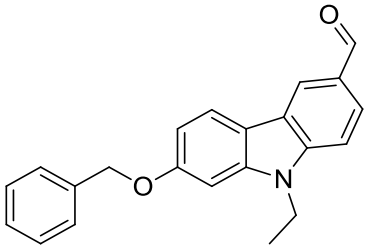
Synthesis of 7-(benzyloxy)-9*H*-carbazole-3-carbaldehyde (**4.23**)



Synthesised according to general procedure D using modified conditions. Reaction was stirred at 165 °C for 18 h. 4'-(Benzyloxy)-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**4.22**) (368 mg, 1.1 mmol), triphenylphosphine (723 mg, 2.75 mmol) and DMA (2 mL). Crude product was purified by chromatography on silica gel using EtOAc/hexane 1/4 to 1/1 as an eluent. Product was isolated as an orange solid, yield 37% (119 mg, 0.39 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 11.69 (s, 1H, NH), 10.01 (s, 1H, CHO), 8.60 (s, 1H, ar CH), 8.12 (d, J = 8.5 Hz, 1H, ar CH), 7.85 (d, J = 8.4 Hz, 1H, ar CH), 7.56 (d, J = 8.4 Hz, 1H, ar CH), 7.50 (d, J = 7.6 Hz, 3H, ar CH), 7.41 (t, J = 7.6 Hz, 3H, ar CH), 7.35 (d, J = 7.1 Hz, 1H, ar CH), 7.12 (s, 1H, ar CH), 6.96 (d, J = 8.5 Hz, 1H, ar CH), 5.22 (s, 2H, CH_2) ppm.

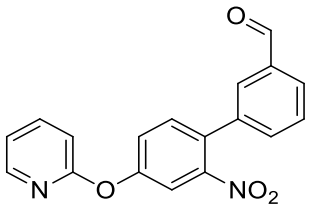
Synthesis of 7-(benzyloxy)-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.24**)



Synthesised according to general procedure E. 7-(Benzyloxy)-9*H*-carbazole-3-carbaldehyde (**4.23**) (119 mg, 0.39 mmol), anhydrous DMF (1 mL), sodium hydride (31 mg, 60% dispersion in mineral oil, 0.78 mmol) and ethyl iodide (62 μL, 0.78 mmol). Crude product was purified by chromatography on silica gel using EtOAc/hexane 1/4 to 1/1 as an eluent. Product was isolated as yellow solid, yield 56% (74 mg, 0.22 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.03 (s, 1H, CHO), 8.62 (s, 1H, ar CH), 8.16 (d, J = 8.5 Hz, 1H, ar CH), 7.91 (d, J = 8.5 Hz, 1H, ar CH), 7.72 (d, J = 8.5 Hz, 1H, ar CH), 7.53 (d, J = 7.6 Hz, 2H, ar CH), 7.42 (t, J = 7.6 Hz, 2H, ar CH), 7.37 – 7.34 (m, 2H, ar CH), 6.98 (d, J = 9.7 Hz, 1H, ar CH), 5.26 (s, 2H, CH_2), 4.47 (q, J = 7.2 Hz, 2H, CH_2), 1.32 (t, J = 7.1 Hz, 3H, CH_3) ppm.

Synthesis of 2'-nitro-4'-(pyridine-2-yloxy)-[1,1'-biphenyl]-3-carbaldehyde (**4.25**)



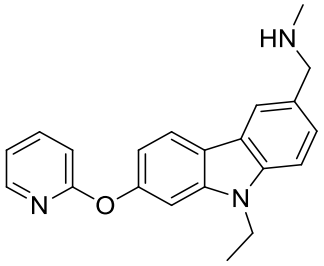
To an oven-dried carosel tube was added 4'-hydroxy-2'-nitro-[1,1'-biphenyl]-3-carbaldehyde (**4.15**) (400 mg, 1.64 mmol), 2-picolinic acid (40 mg, 0.32 mmol), K_3PO_4 (698 mg, 3.28 mmol) and copper (I) iodide (31 mg, 0.16 mmol). The tube was evacuated and backfilled with argon three times. A solution of 2-bromopyridine (188 μL, 1.97 mmol) in DMSO was added and the solution was stirred at 100 °C until TLC indicated completion (3 h). The reaction was cooled and diluted with EtOAc (20 mL) and H_2O (20 mL) and was filtered through celite. Crude product was extracted with EtOAc (3 X 20 mL), the organic extracts were washed with H_2O (20 mL) and dried over anhydrous sodium sulfate. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield crude product that was purified by chromatography on silica gel using EtOAc/hexane 0/1 to 1/4 as an eluent. Product was isolated as a pale yellow solid, yield 64% (341 mg, 1.06 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.05 (s, 1H, CHO), 8.25 – 8.18 (m, 1H, ar CH), 7.95 – 7.89 (m, 1H, ar CH), 7.92 – 7.74 (m, 3H, ar CH), 7.71 – 7.53 (m, 2H, ar CH), 7.56 – 7.41 (m, 2H, ar CH), 7.15 – 7.04 (m, 2H, ar CH) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.7, 162.3, 154.0, 149.0, 147.5, 140.1, 138.4, 136.6, 133.9, 132.8, 131.0, 129.3, 129.3, 129.1, 125.5, 119.7, 117.3, 112.3 ppm.

MS-EI (m/z) found 320 $[M]^+$, calcd 320.

Synthesis of 1-(9-ethyl-7-(pyridine-2-yloxy)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.28**)



Via a three step-synthesis starting from 2'-nitro-4'-(pyridine-2-yloxy)-[1,1'-biphenyl]-3-carbaldehyde (**4.25**). The cyclized carbazole, 1-(9-ethyl-7-(pyridine-2-yloxy)-9*H*-

carbazole-3-carbaldehdye (**4.26**), was synthesised according to general procedure D using modified conditions. Reaction was stirred at 165 °C for 4 h. 2'-Nitro-4'-(pyridine-2-yloxy)-[1,1'-biphenyl]-3-carbaldehyde (**4.25**) (341 mg, 1.06 mmol), triphenyl phosphine (698 mg, 2.66 mmol) and DMA (1 mL). Crude product was purified by chromatography on silica gel using EtOAc/hexane 1/4 to 1/1 as an eluent. Product was isolated as a mixture with starting material (224 mg). Presence of product was confirmed by observation of the NH proton by ^{1}H NMR.

^{1}H NMR (500 MHz, DMSO-d_6) δ 11.84 (s, 1H, NH), 10.22 (d, J = 1.1 Hz, 1H), 10.07 (s, 1H), 8.48 (d, J = 7.6 Hz, 1H), 8.25 – 8.14 (m, 3H), 8.01 – 7.91 (m, 3H), 7.93 – 7.83 (m, 3H), 7.75 – 7.63 (m, 3H), 7.66 – 7.58 (m, 1H), 7.46 – 7.36 (m, 2H), 7.26 – 7.18 (m, 2H), 7.19 – 7.11 (m, 1H), 7.08 – 6.98 (m, 2H). Total 24 x ^{1}H corresponding to both **4.25** and **4.26** in approximately a 1:1 ratio.

9-Ethyl-7-(pyridine-2-yloxy)-9*H*-carbazole-3-carbaldehyde (**4.27**) was synthesised according to general procedure E. A mixture of 1-(9-ethyl-7-(pyridine-2-yloxy)-9*H*-carbazole-3-carbaldehdye (**4.26**) and 2'-nitro-4'-(pyridine-2-yloxy)-[1,1'-biphenyl]-3-carbaldehyde (**4.25**) (114 mg, 0.39 mmol), anhydrous DMF (2 mL), sodium hydride (17 mg, 60% suspension in mineral oil, 0.43 mmol) and ethyl iodide (34 μL, 0.43 mmol). Crude product was purified by chromatography on silica gel using DCM as an eluent. Product was isolated as a mixture where the reaction was confirmed by the presence of CH_2 and CH_3 of the ethyl peak and disappearance of the NH proton.

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.32 (s, 1H), 10.19 (s, 1H), 8.35 – 8.19 (m, 3H), 8.10 (d, J = 8.3 Hz, 2H), 7.92 (dd, J = 7.6, 1.3 Hz, 1H), 7.85 – 7.69 (m, 2H), 7.42 – 7.25 (m, 3H), 7.16 – 6.93 (m, 4H), 4.76 (q, J = 7.1 Hz, 2H, CH_2), 1.44 (t, J = 7.1 Hz, 3H, CH_3).

Synthesis of 1-(9-ethyl-7-(pyridine-2-yloxy)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.28**) was accomplished using general procedure A. 9-Ethyl-7-(pyridine-2-yloxy)-9*H*-carbazole-3-carbaldehyde (**4.27**) (15 mg, 0.047 mmol), methylamine hydrochloride (7 mg, 0.094 mmol), trimethylamine (10 μL, 0.071 mmol), titanium (IV) isopropoxide (28 μL, 0.094 mmol), DCM (1 mL), EtOH (1 mL) and sodium borohydride (4 mg, 0.094 mmol). Product was isolated as a brown oil, yield 66% (10.6 mg, 0.030 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.27 – 8.22 (m, 1H, ar CH), 8.06 (d, J = 8.3 Hz, 1H, ar CH), 8.01 (d, J = 7.7 Hz, 1H, ar CH), 7.73 – 7.66 (m, 1H, ar CH), 7.32 (d, J = 7.3 Hz, 1H, ar CH), 7.21 (s, 1H, ar CH), 7.17 (t, J = 7.5 Hz, 1H, ar CH), 7.04 – 6.98 (m, 2H,

ar CH), 6.92 (d, J = 8.8 Hz, 1H, ar CH), 4.64 (q, J = 7.2 Hz, 2H, CH_2), 4.13 (s, 2H, CH_2), 2.56 (s, 3H, CH_3), 1.43 (t, J = 7.1 Hz, 3H, CH_3) ppm.

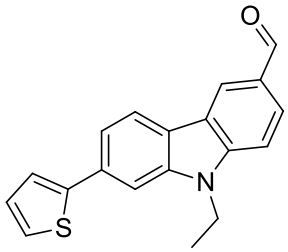
^{13}C NMR (126 MHz, Chloroform-*d*) δ 152.9, 147.9, 141.6, 139.2, 128.3, 120.7, 120.3, 119.7, 118.8, 118.2, 117.7, 115.2, 113.0, 111.1, 109.9, 106.6, 101.7, 65.3, 48.8, 39.3, 15.4 ppm.

HRMS-ESI (m/z) found 301.1331 $[M]^+$, calcd 301.1335 for $[C_{20}H_{17}N_2O]^+$ (loss of NHMe).

LC-MS purity = 90% (UV), ret time = 12.08 min.

General Procedure F

Synthesis of 9-ethyl-7-(thiophen-2-yl)-9*H*-carbazole-3-carbaldehyde (**4.30a**)

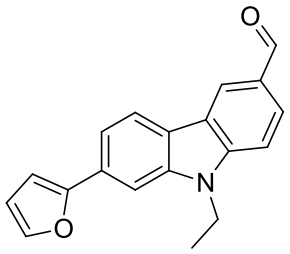


To an oven-dried carosel tube was added 7-bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (500 mg, 1.65 mmol), thiophene-2-boronic acid (254 mg, 1.98 mmol), potassium carbonate (457 mg, 3.30 mmol) and Pd-118 (10 mg, 0.016 mmol). The tube was evacuated and backfilled with argon three times before anhydrous MeCN (2.5 mL) and degassed H_2O (2.5 mL) were added. The solution was heated to 80 °C and stirred for 2 h. The reaction was cooled to room temperature and diluted with EtOAc (30 mL) and H_2O (30 mL). Crude product was extracted with EtOAc (3 X 10 mL), washed with brine (30 mL) and dried over anhydrous $MgSO_4$. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield crude product that was purified by chromatography on silica gel using EtOAc/hexane 1/4 to 1/1 as an eluent (over 15 minutes). Product was isolated as a yellow solid, yield 57 % (288 mg, 0.94 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.08 (s, 1H, CHO), 8.56 (s, 1H, ar CH), 8.10 (d, J = 8.1 Hz, 1H, ar CH), 7.99 (d, J = 8.4 Hz, 1H, ar CH), 7.62 (s, 1H, ar CH), 7.58 (d, J = 8.1 Hz, 1H, ar CH), 7.45 (d, J = 8.6 Hz, 1H, ar CH), 7.43 (d, J = 2.5 Hz, 1H, ar CH), 7.33 (d, J = 4.7 Hz, 1H, ar CH), 7.14 (t, J = 4.2 Hz, 1H, ar CH), 4.40 (q, J = 7.2 Hz, 2H, CH_2), 1.48 (t, J = 7.2 Hz, 3H, CH_3) ppm.

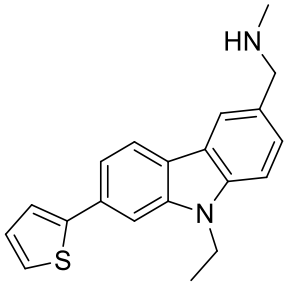
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.6, 144.9, 144.0, 141.1, 133.2, 128.7, 128.1, 127.2, 125.0, 123.8, 123.4, 122.9, 122.4, 121.1, 118.9, 108.7, 106.2, 37.9, 13.8 ppm.
HRMS-ESI (*m/z*) found 306.0947 $[M+H]^+$, cald 306.0947 for $[C_{19}H_{15}NOS+H]^+$.

Synthesis of 9-ethyl-7-(furan-2-yl)-9H-carbazole-3-carbaldehyde (**4.30b**)



Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (500 mg, 1.65 mmol), furan-2-boronic acid (222 mg, n1.98 mmol), potassium carbonate (457 mg, 3.30 mmol), Pd-118 (10 mg, 0.16 mmol), anhydrous MeCN (2.5 mL) and degassed H_2O (2.5 mL). Product was isolated as a yellow solid, yield 56% (271 mg, 0.93 mmol).
^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.10 (s, 1H, CHO), 8.58 (s, 1H, ar CH), 8.13 (d, *J* = 8.5 Hz, 1H, ar CH), 8.01 (d, *J* = 8.5 Hz, 1H, ar CH), 7.77 (s, 1H, ar CH), 7.63 (d, *J* = 8.2 Hz, 1H, ar CH), 7.54 (s, 1H, ar CH), 7.47 (d, *J* = 8.5 Hz, 1H, ar CH), 6.79 (d, *J* = 3.1 Hz, 1H, ar CH), 6.57 – 6.48 (m, 1H, ar CH), 4.44 (q, *J* = 7.2 Hz, 2H, CH_2), 1.51 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.
^{13}C NMR (126 MHz, DMSO-d_6) δ 192.2, 154.1, 144.1, 143.3, 141.2, 129.4, 128.9, 127.1, 124.3, 122.6, 122.0, 121.8, 116.6, 112.7, 110.0, 106.7, 104.8, 37.8, 14.2 ppm.
HRMS-ESI (m/z) found 312.0991 $[M+Na]^+$, calcd 312.0995 for $[C_{19}H_{15}NO_2+Na]^+$.

Synthesis of 1-(9-ethyl-7-(thiophen-2-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31a**)



Synthesised according to general procedure B. 9-Ethyl-7-(thiophen-2-yl)-9*H*-carbazole-3-carbaldehyde (**4.30a**) (100 mg, 0.32 mmol), methylamine (163 µL, 2M solution in THF,

0.32 mmol), anhydrous THF (5 mL) and sodium triacetoxyborohydride (104 mg, 0.49 mmol). Product isolated as a pale yellow solid, yield 12 % (13 mg, 0.040 mmol).

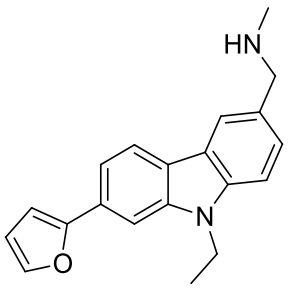
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.05 (d, *J* = 8.3 Hz, 1H, ar CH), 8.04 (s, 1H, ar CH), 7.58 (s, 1H, ar CH), 7.51 – 7.48 (m, 1H, ar CH), 7.46 – 7.42 (m, 1H, ar CH), 7.41 (d, *J* = 3.6 Hz, 1H, ar CH), 7.35 (d, *J* = 8.3 Hz, 1H, ar CH), 7.32 – 7.27 (m, 1H, ar CH), 7.13 (d, *J* = 3.6 Hz, 1H, ar CH), 7.12 (d, *J* = 3.6 Hz, 1H, ar CH), 4.37 (q, *J* = 7.2 Hz, 2H, CH_2), 3.95 (s, 2H, CH_2), 2.53 (s, 2H, CH_2), 1.44 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 145.6, 140.6, 139.9, 132.1, 128.4, 127.9, 126.4, 124.4, 123.0, 122.8, 122.3, 120.7, 120.3, 117.6, 108.4, 105.7, 55.9, 37.5, 35.3, 13.7 ppm.

HRMS-ESI (*m/z*) found 290.0992 $[M]^+$, calcd 290.0998 $[C_{19}H_{16}NS]^+$ (loss of NHMe).

LC-MS purity > 99% (UV), ret. time = 12.73 min.

Synthesis of 1-(9-ethyl-7-(furan-2-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31b**)



Synthesised according to general procedure B. 9-Ethyl-7-(furan-2-yl)-9H-carbazole-3-carbaldehyde (**4.30b**) (100 mg, 0.34 mmol), methylamine (259 µL, 2 M solution in THF, 0.51 mmol), anhydrous THF (4 mL) and sodium triacetoxyborohydride (183 mg, 0.86 mmol). Product isolated as an orange solid, yield 46% (48 mg, 0.15 mmol).

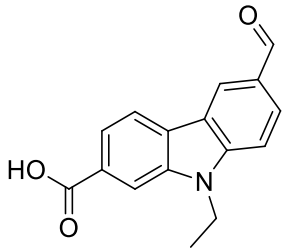
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.06 (d, *J* = 8.1 Hz, 1H, ar CH), 8.03 (s, 1H, ar CH), 7.72 (s, 1H, ar CH), 7.54 (d, *J* = 8.2 Hz, 1H, ar CH), 7.53 – 7.50 (m, 1H, ar CH), 7.42 (d, *J* = 8.1 Hz, 1H, ar CH), 7.35 (d, *J* = 8.2 Hz, 1H, ar CH), 6.75 (d, *J* = 3.3 Hz, 1H, ar CH), 6.55 – 6.51 (m, 1H, ar CH), 4.38 (q, *J* = 7.2 Hz, 2H, CH_2), 3.93 (s, 2H, CH_2), 2.53 (s, 3H, CH_3) and 1.45 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 155.0, 141.7, 140.5, 139.8, 130.4, 128.5, 126.3, 122.9, 122.1, 120.6, 120.1, 115.4, 111.7, 108.3, 104.8, 103.6, 56.2, 37.5, 35.7 and 13.8 ppm.

HRMS-ESI (*m/z*) found 274.1219 $[M]^+$, calcd 274.1226 $[C_{19}H_{16}NO]^+$ (loss of NHMe).

LC-MS purity > 99% (UV), ret. time = 12.38 min.

Synthesis of 9-ethyl-6-formyl-9*H*-carbazole-2-carboxylic acid (**4.33**)



To a solution of methyl-9-ethyl-6-formyl-9*H*-carbazole-2-carboxylate (**4.13**) (138 mg, 0.49 mmol) in MeOH (10 mL) and H_2O (10 mL) was added lithium hydroxide (58 mg, 2.50 mmol) and the resulting solution was heated at 90 °C for 2 h. The solution was cooled and diluted with H_2O (10 mL) and neutralized with 1 N HCl. Crude product was extracted with EtOAc (3 X 10 mL) and dried over anhydrous $MgSO_4$. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield product as a white solid, yield 80% (106 mg, 0.40 mmol).

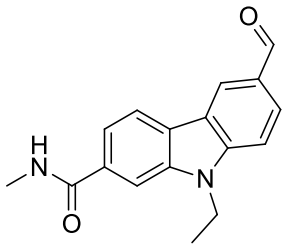
^{1}H NMR (500 MHz, DMSO-d_6) δ 13.06 (s, 1H, CO_2H), 10.06 (s, 1H, CHO), 8.84 (s, 1H, ar CH), 8.39 (d, J = 8.1 Hz, 1H, ar CH), 8.25 (s, 1H, ar CH), 8.05 (d, J = 8.6 Hz, 1H, ar CH), 7.88 (d, J = 8.1 Hz, 1H, ar CH), 7.85 (d, J = 8.6 Hz, 1H, ar CH), 4.58 (q, J = 7.1 Hz, 2H, CH_2), 1.34 (t, J = 7.1 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.2, 168.1, 140.2, 129.1, 127.9, 126.3, 125.5, 122.0, 121.4, 121.2, 113.8, 111.6, 110.5, 109.9, 37.9, 14.2 ppm.

MS-EI (m/z) found 266 $[M]^-$, calcd 266.

General Procedure G

Synthesis of 9-ethyl-6-formyl-*N*-methyl-9*H*-carbazole-2-carboxamide (**4.33a**)



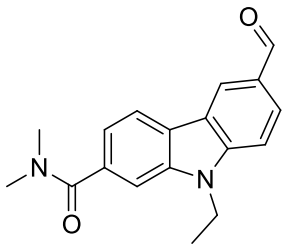
To an oven-dried two-neck flask was added 9-ethyl-6-formyl-9*H*-carbazole-2-carboxylic acid (**4.33**) (30 mg, 0.11 mmol), HATU (42 mg, 0.11 mmol) and the flask was evacuated and backfilled with argon 3 times. Anhydrous DMF (1 mL) and trimethylamine (31 μL, 0.22 mmol) were added and the solution was stirred at rt for 5 minutes. Methylamine

hydrochloride was added and the solution was stirred until TLC indicated completion (18 h). The solution was diluted with EtOAc (30 mL) and H_2O (30 mL) and crude product was extracted with EtOAc (3 X 30 mL). The organic extracts were washed with saturated $NaHCO_3$, dried over anhydrous $MgSO_4$, filtered and the filtrate was concentrated under a reduced pressure to yield crude product. Product was purified by chromatography on silica gel using DCM/MeOH 9/1 as an eluent to yield product as a yellow solid, yield 90% (27 mg, 0.096 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.12 (s, 1H, CHO), 8.64 (s, 1H, ar CH), 8.17 (d, *J* = 8.1 Hz, 1H, ar CH), 8.11 – 8.03 (m, 2H, ar CH), 7.59 (d, *J* = 8.1 Hz, 1H, ar CH), 7.53 (d, *J* = 8.5 Hz, 1H, ar CH), 6.33 (br s, 1H, NH), 4.47 (q, *J* = 7.3 Hz, 2H, CH_2), 3.11 (s, 3H, CH_3), 1.50 (t, *J* = 7.3 Hz, 3H, CH_3) ppm.

HRMS-ESI (m/z) found 281.1285 $[M+H]^+$, calcd 281.1245 for $[C_{17}H_{16}N_2O_2+H]^+$.

Synthesis of 9-ethyl-6-formyl-*N,N*-dimethyl-9*H*-carbazole-2-carboxamide (**4.33b**)

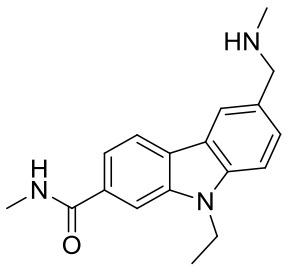


Synthesised according to general procedure G. 9-Ethyl-6-formyl-9*H*-carbazole-2-carboxylic acid (**4.33**) (30 mg, 0.11 mmol), HATU (42 mg, 0.11 mmol), anhydrous DMF (1 mL), triethylamine (31 μL, 0.22 mmol) and dimethylamine hydrochloride (11 mg, 0.13 mmol). Product was isolated as an orange oil, yield 84% (28 mg, 0.095 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.09 (s, 1H, CHO), 8.61 (s, 1H, ar CH), 8.14 (d, *J* = 7.9 Hz, 1H, ar CH), 8.03 (d, *J* = 8.5 Hz, 1H, ar CH), 7.59 (s, 1H, ar CH), 7.50 (d, *J* = 8.5 Hz, 1H, ar CH), 7.35 (d, *J* = 7.9 Hz, 1H, ar CH), 4.41 (q, *J* = 7.2 Hz, 2H, CH_2), 3.12 (s, 6H, CH_3), 1.47 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

HRMS-ESI (m/z) found 317.1265 $[M+Na]^+$, calcd 317.1260 for $[C_{18}H_{18}N_2O_2+Na]^+$.

Synthesis of 9-ethyl-*N*-methyl-6-((methylamino)methyl)-9*H*-carbazole-2-carboxamide (**4.34a**)



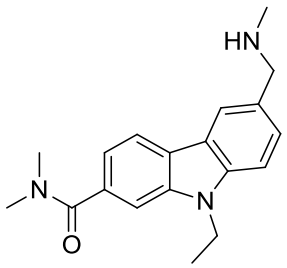
Synthesised according to general procedure A. 9-Ethyl-6-formyl-*N*-methyl-9*H*-carbazole-2-carboxamide (**4.33a**) (12 mg, 0.04 mmol), methylamine hydrochloride (6 mg, 0.08 mmol), triethylamine (9 µL, 0.06 mmol), titanium (IV) isopropoxide (25 µL, 0.08 mmol), DCM (2 mL), EtOH, (2 mL) and sodium borohydride (3 mg, 0.08 mmol). Product was isolated as a yellow solid, yield 27% (3 mg, 0.010 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.11 – 8.03 (m, 2H, ar CH), 7.98 (s, 1H, ar CH), 7.48 (d, *J* = 8.2 Hz, 2H, ar CH), 7.39 (d, *J* = 8.3 Hz, 1H, ar CH), 6.37 (br s, H, NH), 4.39 (q, *J* = 7.2 Hz, 2H, CH_2), 3.94 (s, 2H, CH_2), 3.09 (d, *J* = 4.7 Hz, 3H, CH_3), 2.53 (s, 3H, CH_3), 1.43 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

MS-ESI (m/z) found 265.2 $[M]^+$ (LC-MS), calcd 265.1 for $[C_{17}H_{17}N_2O]^+$ (loss of NHMe).

LC-MS purity > 99% (UV), ret time = 3.02 min.

Synthesis of 9-ethyl-*N*,*N*-dimethyl-6-((methylamino)methyl)-9*H*-carbazole-2-carboxamide (**4.34b**)



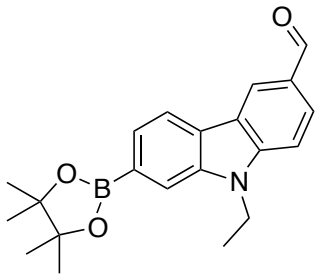
Synthesised according to general procedure A. 9-Ethyl-6-formyl-*N*,*N*-dimethyl-9*H*-carbazole-2-carboxamide (**4.33b**) (28 mg, 0.09 mmol), methylamine hydrochloride (13 mg, 0.19 mmol), triethylamine (18 µL, 0.135 mmol), titanium (IV) isopropoxide (56 µL, 0.19 mmol), DCM (1 mL), EtOH (1 mL) and sodium borohydride (7 mg, 0.19 mmol). Product was isolated as an orange oil, yield 14% (4 mg, 0.012 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.12 – 8.03 (m, 2H, ar CH), 7.52 (s, 1H, ar CH), 7.47 (d, *J* = 8.3 Hz, 1H, ar CH), 7.38 (d, *J* = 8.3 Hz, 1H, ar CH), 7.25 (d, *J* = 7.9 Hz, 1H, ar CH), 4.37 (q, *J* = 7.1 Hz, 2H, CH_2), 3.94 (s, 2H, CH_2), 3.12 (br s, 6H, CH_3), 2.52 (s, 2H, CH_2), 1.68 (s, 3H, CH_3), 1.44 (t, *J* = 7.1 Hz, 3H, CH_3) ppm.

MS-ESI (m/z) found 279.2 $[M]^+$ (LC-MS), calcd 272.1 for $[C_{18}H_{19}N_2O]^+$ (loss of NHMe).

LC-MS purity > 99% (UV), ret time = 2.84 min.

Synthesis of 9-ethyl-7-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-9H-carbazole-3-carbaldehyde (**4.37**)



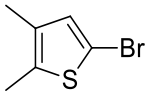
To a solution of 7-bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (900 mg, 2.9 mmol), bispinacolato diboron (1.134 g, 4.4 mmol) and KOAc (876 mg, 8.9 mmol) in 1,4-dioxane (10 mL) under argon was added $Pd(dppf)Cl_2.CH_2Cl_2$ (243 mg, 0.29 mmol). The resulting solution was stirred at 80 °C for 2 hours. The solution was cooled and H_2O was added (50 mL). Crude product was extracted with EtOAc (3 X 50 mL). The organic extracts were washed with brine (50 mL) then dried over anhydrous $MgSO_4$. The suspension was filtered and the filtrate was concentrated under a reduced pressure to yield crude product which was purified by chromatography on silica gel (hexane/EtOAc 4/1 as an eluent). Product isolated as a pale yellow solid, yield 95% (966 mg, 2.76 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.04 (s, 1H, CHO), 8.76 (s, 1H, ar CH), 8.28 (d, *J* = 7.8 Hz, 1H, ar CH), 8.00 (d, *J* = 8.5 Hz, 1H, ar CH), 7.91 (s, 1H, ar CH), 7.79 (d, *J* = 8.6 Hz, 1H, ar CH), 7.61 (d, *J* = 7.8 Hz, 1H, ar CH), 4.53 (q, J = 6.9 Hz, 2H, CH_2), 1.33 (m, 15H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.2, 143.9, 140.2, 128.8, 127.5, 126.3, 125.4, 125.1, 122.4, 120.8, 116.0, 110.3, 84.2 (2C), 37.8, 25.2 (4C), 14.2 ppm.

HRMS-ESI (m/z) found 372.1745 $[M+Na]^+$, calcd 372.1741 for $[C_{21}H_{24}BNO_3+Na]^+$.

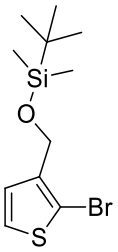
Synthesis of 5-bromo-2,3-dimethylthiophene (**4.38**)



Synthesised according to a literature procedure.[533] To a solution of 2,3-dimethylthiophene (500 µL, 4.4 mmol) in DCM (10 mL) was added N-bromosuccinimide (872 mg, 4.9 mmol). The resulting solution was stirred at room temperature for 18 h. The solution was diluted with DCM (10 mL) and silica was added. Solvent was removed under a reduced pressure to yield crude product loaded on silica. Crude product was purified by chromatography on silica gel using hexane as an eluent to yield product as a light brown oil in a quantitative yield (840 mg, 4.39 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 6.75 (s, 1H, ar CH), 2.31 (s, 3H, CH_3), 2.13 (s, 3H, CH_3) ppm. Consistent with the literature values.[534]

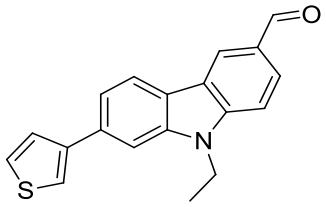
Synthesis of ((2-bromothiophen-3-yl)methoxy)(tert-butyl)dimethylsilane (**4.39**)



To a solution of ((2-bromo-thiophen-3-yl)methoxy)(*tert*-butyl)dimethylsilane (533 mg, 2.76 mmol) and *tert*-butyldimethylsilyl chloride (457 mg, 3.03 mmol) in anhydrous DCM (5 mL) was added trimethylamine (577 µL, 4.14 mmol) at 0 °C. The solution was allowed to warm to room temperature and was stirred until TLC indicated completion (18 h). The suspension was poured into saturated sodium hydrogen carbonate and DCM was added. The solution was filtered through a hydrophobic frit and the organic extract was filtered through a silica plug, eluting product with DCM. Solvent was removed under a reduced pressure to yield product as a colourless liquid yield 53% (454 mg, 1.47 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.23 (d, *J* = 5.6 Hz, 1H, ar CH), 7.04 (d, *J* = 5.6 Hz, 1H, ar CH), 4.67 (s, 2H, CH_2), 0.97 (s, 9H, CH_3), 0.14 (s, 6H, CH_3) ppm. Consistent with the literature values.

Synthesis of 9-ethyl-7(thiophen-3-yl)-9*H*-carbazole-3-carbaldehyde (**4.30c**)



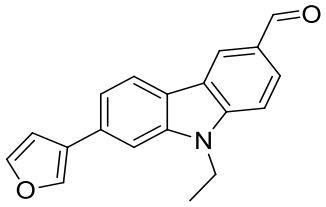
Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (500 mg, 1.65 mmol), thiophene-3-boronic acid (254 mg, 1.98 mmol), potassium carbonate (457 mg, 3.30 mmol), Pd-118 (10 mg, 0.016 mmol), anhydrous MeCN (2.5 mL) and degassed H_2O (2.5 mL). Product was isolated as a orange solid, yield 74% (374 mg, 1.22 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.10 (s, 1H, CHO), 8.60 (s, 1H, ar CH), 8.15 (d, *J* = 8.0 Hz, 1H, ar CH), 8.01 (d, *J* = 8.5 Hz, 1H, ar CH), 7.63 (s, 1H, ar CH), 7.61 – 7.55 (m, 2H, ar CH), 7.51 (d, *J* = 4.8 Hz, 1H, ar CH), 7.48 (d, *J* = 8.5 Hz, 1H, ar CH), 7.45 (d, *J* = 2.9 Hz, 1H, ar CH), 7.45 (d, *J* = 3.0 Hz, 1H, ar CH), 4.44 (q, *J* = 7.2 Hz, 2H, CH_2), 1.50 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.5, 144.0, 142.8, 141.2, 134.9, 128.8, 127.1, 126.6, 126.3, 123.7, 123.1, 122.1, 121.0, 120.6, 119.4, 108.6, 106.8, 37.9, 13.8 ppm.

HRMS-ESI (*m/z*) found 328.0768 $[M+Na]^+$, calcd 328.0767 $[C_{19}H_{15}NOS+Na]^+$.

Synthesis of 9-ethyl-7-(furan-3-yl)-9*H*-carbazole-3-carbaldehyde (**4.30d**)

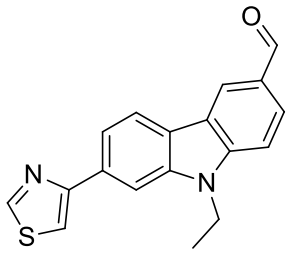


Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (500 mg, 1.65 mmol), furan-3-boronic acid (221 mg, 1.98 mmol), potassium carbonate (457 mg, 3.30 mmol), Pd-118 (10 mg, 0.016 mmol), anhydrous MeCN (2.5 mL) and degassed H_2O (2.5 mL). Product was isolated as a pale yellow solid, yield 59% (283 mg, 0.97 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.08 (s, 1H, CHO), 8.57 (s, 1H, ar CH), 8.11 (d, *J* = 8.0 Hz, 1H, ar CH), 7.99 (d, *J* = 8.5 Hz, 1H, ar CH), 7.85 (s, 1H, ar CH), 7.54 (d, *J* = 1.6 Hz, 1H, ar CH), 7.50 (s, 1H, ar CH), 7.48 – 7.42 (m, 2H, ar CH), 6.85 – 6.81 (m, 1H, ar CH), 4.41 (q, *J* = 7.2 Hz, 2H, CH_2), 1.48 (t, *J* = 7.3 Hz, 3H, CH_3) ppm.

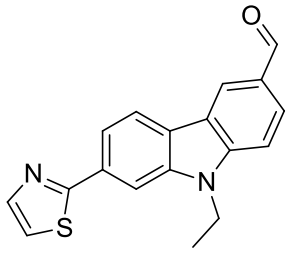
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.7, 143.9, 143.8, 141.2, 138.7, 131.3, 128.6, 127.1, 127.0, 123.7, 123.0, 122.0, 121.1, 118.7, 109.1, 108.6, 106.1, 37.9, 13.8 ppm.
HRMS-ESI (*m/z*) found 312.0991 $[M+Na]^+$, calcd 312.0995 for $[C_{19}H_{15}NO_2+Na]^+$.

Synthesis of 9-ethyl-7-(thiazol-4-yl)-9H-carbazole-3-carbaldehyde (**4.30e**)



Synthesised according to general procedure F. 9-Ethyl-7-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-9H-carbazole-3-carbaldehyde (**4.37**) (470 mg, 1.34 mmol), 4-bromothiazole (100 μL, 1.12 mmol), potassium carbonate (308 mg, 2.23 mmol), Pd-118 (36 mg, 0.05 mmol), anhydrous MeCN (5 mL) and degassed H_2O (5 mL). Product was isolated as a brown oil, yield 99% (408 mg, 1.33 mmol).
^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.03 (s, 1H), 8.90 (d, *J* = 1.6 Hz, 1H), 8.50 (s, 1H), 8.08 – 8.04 (m, 2H), 7.94 (d, *J* = 8.5 Hz, 1H), 7.73 (d, *J* = 8.8 Hz, 1H), 7.60 (d, *J* = 1.8 Hz, 1H), 7.38 (d, *J* = 8.5 Hz, 1H), 4.35 (q, *J* = 7.3 Hz, 2H), 1.42 (t, *J* = 7.3 Hz, 3H) ppm.
HRMS-ESI (m/z) found 307.0889 $[M+H]^+$, calcd 307.0860 for $[C_{18}H_{14}N_2OS+H]^+$.

Synthesis of 9-ethyl-7-(thiazol-2-yl)-9H-carbazole-3-carbaldehyde (**4.30f**)

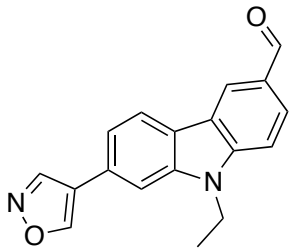


Synthesised according to general procedure F. 9-Ethyl-7-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-9H-carbazole-3-carbaldehyde (**4.37**) (500 mg, 1.4 mmol), 2-bromothiazole (86 μL, 0.95 mmol), potassium carbonate (262 mg, 1.90 mmol), Pd-118 (31 mg, 0.004 mmol), anhydrous MeCN (4 mL) and degassed H_2O (4 mL). Product was isolated as a orange solid, yield 30% (132 mg, 0.43 mmol).
^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.10 (s, 1H, CHO), 8.60 (s, 1H, ar CH), 8.16 (s, 2H, ar CH), 8.03 (d, *J* = 8.4 Hz, 1H, ar CH), 7.92 (s, 1H, ar CH), 7.85 (d, *J* = 8.1 Hz, 1H,

ar CH), 7.49 (d, *J* = 8.5 Hz, 1H, ar CH), 7.38 (s, 1H, ar CH), 4.46 (q, *J* = 7.2 Hz, 2H, CH_2), 1.50 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

HRMS-ESI (m/z) found 307.0900 $[M+H]^+$, calcd 307.0900 for $[C_{18}H_{14}N_2OS+H]^+$.

Synthesis of 9-ethyl-7-(isoxazol-4-yl)-9*H*-carbazole-3-carbaldehyde (**4.30g**)



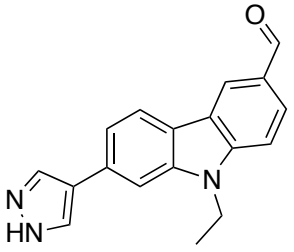
Synthesised according to general procedure F under anhydrous conditions, using potassium phosphate tribasic as a base. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (107 mg, 0.35 mmol), isoxazole-4-boronic acid (120 mg, 1.06 mmol) K_3PO_4 (300 mg, 1.41 mmol), Pd-118 (11 mg, 0.01 mmol), anhydrous MeCN (2 mL). Product was isolated as an orange solid, yield 31% (32 mg, 0.11 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.05 (s, 1H, CHO), 9.55 (s, 1H, ar CH), 9.31 (s, 1H, ar CH), 8.75 (s, 1H, ar CH), 8.33 (d, *J* = 8.1 Hz, 1H, ar CH), 8.08 (s, 1H, ar CH), 7.99 (d, *J* = 9.9 Hz, 1H, ar CH), 7.78 (d, *J* = 8.5 Hz, 1H, ar CH), 7.64 (d, *J* = 8.1 Hz, 1H, ar CH), 4.54 (q, *J* = 7.2 Hz, 2H, CH_2), 1.38 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 192.2, 155.5, 148.9, 144.0, 141.4, 129.0, 127.5, 127.2, 124.3, 122.7, 122.3, 122.0, 121.9, 119.1, 110.1, 107.9, 37.9, 14.2 ppm.

HRMS-ESI (*m/z*) found 291.1124 $[M+H]^+$, calcd for $[C_{18}H_{14}N_2O_2+H]^+$.

Synthesis of 9-ethyl-7-(1*H*-pyrazol-4-yl)-9*H*-carbazole-3-carbaldehyde (**4.30h**)



Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (500 mg, 1.65 mmol), tert-butyl 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1H-pyrazole-1-carboxylate (584 mg, 1.98 mmol), potassium

carbonate (457 mg, 3.30 mmol), Pd-118 (10 mg, 0.016 mmol), anhydrous MeCN (2.5 mL) and degassed H_2O (2.5 mL). The below data refer to the Boc-protected product.

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.08 (s, 1H, CHO), 8.58 (s, 1H, ar CH), 8.42 (s, 1H, ar CH), 8.14 (d, *J* = 8.0 Hz, 1H, ar CH), 8.11 (s, 1H, ar CH), 8.00 (d, *J* = 8.4 Hz, 1H, ar CH), 7.54 (s, 1H, ar CH), 7.52 – 7.41 (m, 2H, ar CH), 4.42 (q, *J* = 6.9 Hz, 2H, CH_2), 1.70 (s, 9H, CH_3), 1.49 (t, *J* = 6.9 Hz, 3H, CH_3) ppm.

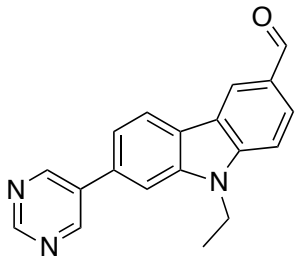
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.7, 144.0, 142.0, 141.1, 129.5, 128.7, 127.3, 126.6, 126.0, 123.8, 122.9, 122.4, 121.3, 118.7, 108.7, 106.2, 85.8, 37.9, 27.9 (3C), 24.8, 13.8 ppm.

HRMS-ESI (m/z) found 412.1650 $[M+Na]^+$, calcd 412.1632 for $[C_{23}H_{23}N_3O_3+Na]^+$.

Boc- deprotection using 4 N HCl in dioxane (2 mL) yielded product as a white solid, where successful deprotection was confirmed by ^{1}H NMR, yield 16% (103 mg, 0.26 mmol).

^{1}H NMR (500 MHz, DMSO-d_6) δ 12.95 (s, 1H, NH), 10.04 (s, 1H, CHO), 8.69 (s, 1H, ar CH), 8.33 (s, 1H, ar CH), 8.22 (d, *J* = 8.1 Hz, 1H, ar CH), 8.08 (s, 1H, ar CH), 7.98 – 7.89 (m, 2H, ar CH), 7.74 (d, *J* = 8.5 Hz, 1H, ar CH), 7.57 (d, *J* = 8.1 Hz, 1H, ar CH), 4.53 (q, *J* = 7.0 Hz, 2H, CH_2), 1.36 (t, *J* = 7.0 Hz, 3H, CH_3) ppm.

Synthesis of 9-ethyl-7-(pyrimidin-5-yl)-9*H*-carbazole-3-carbaldehyde (**4.30i**)

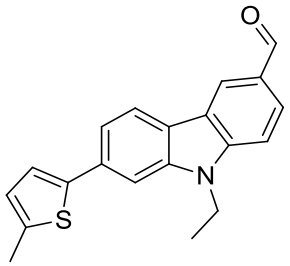


Synthesised according to general procedure F under anhydrous conditions, using potassium phosphate tribasic as a base. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (97 mg, 0.32 mmol), pyrimidine-5-boronic acid (120 mg, 0.96 mmol), K_3PO_4 (272 mg, 1.28 mmol), Pd-118 (10 mg, 0.01 mmol), anhydrous MeCN (2 mL). Product isolated as an orange solid, yield 46% (45 mg, 0.149 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.11 (s, 1H, CHO), 9.24 (s, 1H, ar CH), 9.07 (s, 2H, ar CH), 8.64 (s, 1H, ar CH), 8.27 (d, *J* = 8.0 Hz, 1H, ar CH), 8.05 (d, *J* = 8.5 Hz, 1H, ar CH), 7.61 (s, 1H, ar CH), 7.52 (d, *J* = 8.6 Hz, 1H, ar CH), 4.47 (q, *J* = 7.1 Hz, 2H, CH_2), 1.52 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

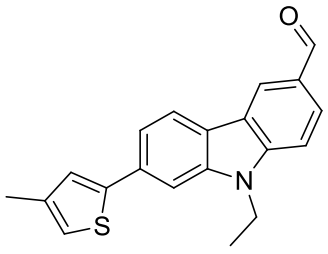
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.4, 157.4, 155.1 (2C), 144.1, 141.2, 134.9, 132.9, 129.1, 127.8, 124.1, 123.6, 122.6, 121.7, 119.3, 108.7, 107.5, 38.0, 13.8 ppm.
HRMS-ESI (*m/z*) found 302.1286 $[M+H]^+$, calcd 302.1288 for $[C_{19}H_{15}N_3O+H]^+$.

Synthesis of 9-ethyl-7-(5-methylthiophen-2-yl)-9H-carbazole-3-carbaldehyde (**4.30j**)



Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (180 mg, 0.59 mmol), (5-methylthiophen-2-yl)boronic acid (254 mg, 1.79), potassium carbonate (164 mg, 1.19 mmol), Pd-118 (4 mg, 0.005 mmol), anhydrous MeCN (2 mL) and degassed H_2O (2 mL). Product was isolated as a brown solid, yield 38% (73 mg, 0.22 mmol).
^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.08 (s, 1H, CHO), 8.55 (s, 1H, ar CH), 8.07 (d, *J* = 8.0 Hz, 1H, ar CH), 7.98 (d, *J* = 8.4 Hz, 1H, ar CH), 7.58 – 7.49 (m, 2H, ar CH), 7.44 (d, *J* = 8.4 Hz, 1H, ar CH), 7.23 (s, 1H, ar CH), 6.82 – 6.73 (m, 1H, ar CH), 4.39 (q, *J* = 6.9 Hz, 2H, CH_2), 2.55 (s, 3H, CH_3), 1.47 (t, *J* = 7.1 Hz, 3H, CH_3) ppm.
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.7, 144.0, 142.4, 141.1, 139.8, 133.5, 128.7, 127.1, 126.3, 123.7, 123.3, 123.0, 122.0, 121.0, 118.5, 108.6, 105.7, 37.9, 15.5, 13.8 ppm.
HRMS-ESI (m/z) found 342.0917 $[M+Na]^+$, calcd 342.0923 for $[C_{20}H_{17}NOS+Na]^+$.

Synthesis of 9-Ethyl-7-(4-methylthiophene-2-yl)-9*H*-carbazole-3-carbaldehyde (**4.30k**)



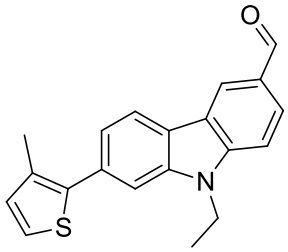
Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (300 mg, 0.99 mmol), (4-methylthiophen-2-yl)boronic acid (422 mg, 2.97 mmol), potassium carbonate (274 mg, 1.98 mmol), Pd-118 (6 mg, 0.009 mmol),

anhydrous MeCN (5 mL) and degassed H_2O (5 mL). Product was isolated as a brown solid, yield 90% (285 mg, 0.89 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.09 (s, 1H, CHO), 8.59 (s, 1H, ar CH), 8.12 (d, *J* = 8.0 Hz, 1H, ar CH), 8.01 (d, *J* = 8.5 Hz, 1H, ar CH), 7.61 (s, 1H, ar CH), 7.57 (d, *J* = 8.0 Hz, 1H, ar CH), 7.48 (d, *J* = 8.5 Hz, 1H, ar CH), 6.91 (s, 1H, ar CH), 4.44 (q, *J* = 7.0 Hz, 2H, CH_2), 2.33 (s, 3H, CH_3), 1.50 (t, *J* = 7.0 Hz, 3H, CH_3) ppm.

HRMS-ESI (m/z) found 342.0932 $[M+Na]^+$, calc 342.0923 for $[C_{20}H_{17}NOS+Na]^+$.

Synthesis of 9-ethyl-7-(3-methylthiophen-2-yl)-9H-carbazole-3-carbaldehyde (**4.30l**)



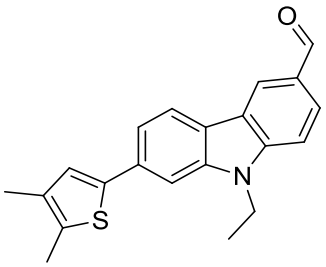
Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (150 mg, 0.49 mmol), (3-methylthiophen-2-yl)boronic acid (211 mg, 1.48 mmol), potassium carbonate (137 mg, 0.99 mmol), Pd-118 (4 mg, 0.004 mmol), anhydrous MeCN (2 mL) and degassed H_2O (2 mL). Product was isolated as a brown solid, yield 75% (120 mg, 0.37 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.10 (s, 1H, CHO), 8.61 (s, 1H, ar CH), 8.16 (d, *J* = 8.0 Hz, 1H, ar CH), 8.02 (d, *J* = 8.5 Hz, 1H, ar CH), 7.52 (s, 1H, ar CH), 7.49 (d, *J* = 8.5 Hz, 1H, ar CH), 7.44 (d, *J* = 8.0 Hz, 1H, ar CH), 7.29 – 7.23 (m, 2H, ar CH), 6.99 (d, *J* = 5.1 Hz, 1H, ar CH), 4.42 (q, *J* = 7.1 Hz, 2H, CH_2), 2.41 (s, 3H, CH_3), 1.49 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.7, 144.0, 140.8, 138.3, 133.4, 133.3, 131.2, 128.7, 127.3, 123.9, 123.7, 122.9, 122.1, 121.7, 120.7, 109.5, 108.7, 38.0, 15.0, 13.8 ppm.

HRMS-ESI (m/z) found 320.1091 $[M+H]^+$, calcd 320.1104 for $[C_{20}H_{17}NOS+H]^+$.

Synthesis of 7-(4,5-dimethylthiophen-2-yl)-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.30m**)



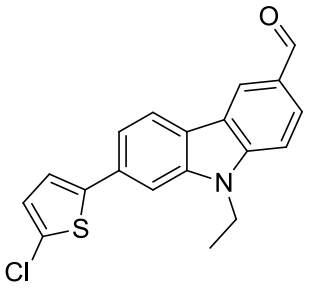
Synthesised according to general procedure F. 5-Bromo-2,3-dimethylthiophene (**4.38**) (121 mg, 0.63 mmol), 9-ethyl-7-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-9H-carbazole-3-carbaldehyde (**4.37**) (332 mg, 0.95 mmol), potassium carbonate (174 mg, 1.26 mmol), Pd-118 (4 mg, 1.26 mmol), anhydrous MeCN (1 mL) and degassed H_2O (1 mL). Product was isolated as a brown solid, yield 46 % (153 mg, 0.45 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.07 (s, 1H, CHO), 8.53 (s, 1H, ar CH), 8.06 (d, *J* = 8.0 Hz, 1H, ar CH), 7.97 (d, *J* = 8.5 Hz, 1H, ar CH), 7.56 – 7.47 (m, 2H, ar CH), 7.43 (d, *J* = 9.2 Hz, 1H, ar CH), 7.13 (s, 1H, ar CH), 4.37 (q, *J* = 7.0 Hz, 2H, CH_2), 2.40 (s, 3H, CH_3), 2.19 (s, 3H, CH_3), 1.46 (t, *J* = 6.6 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.7, 144.0, 141.1, 139.6, 134.3, 133.5, 132.9, 128.6, 127.1, 126.3, 123.7, 123.0, 121.9, 121.0, 118.4, 108.6, 105.5, 37.8, 13.8 (2C), 13.2 ppm.

HRMS-ESI (m/z) found 334.1262 $[M+H]^+$, calcd 334.1260 for $[C_{21}H_{19}NOS+H]^+$.

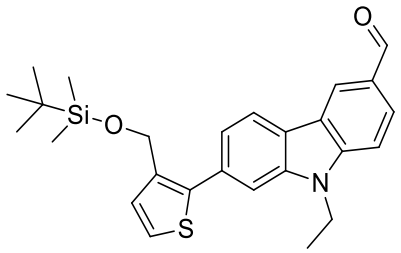
Synthesis of 7-(5-chlorothiophen-2-yl)-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.30n**)



Synthesised according to general procedure F. 7-Bromo-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.4**) (150 mg, 0.49 mmol), (5-chlorothiophen-2-yl)boronic acid (241 mg, 1.48 mmol), potassium carbonate (137 mg, 0.99 mmol), Pd-118 (4 mg, 0.004 mmol), anhydrous MeCN (2 mL) and degassed H_2O (2 mL). Product was isolated as a yellow solid, yield 88% (96 mg, 0.28 mmol).

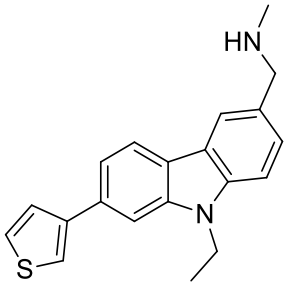
^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.06 (s, 1H, CHO), 8.52 (s, 1H, ar CH), 8.04 (d, *J* = 8.0 Hz, 1H, ar CH), 7.97 (d, *J* = 8.4 Hz, 1H, ar CH), 7.46 (s, 1H, ar CH), 7.44 – 7.37 (m, 2H, ar CH), 7.18 – 7.13 (m, 1H, ar CH), 6.95 – 6.87 (m, 1H, ar CH), 4.34 (q, *J* = 7.0 Hz, 2H, CH_2), 1.45 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.6, 144.0, 143.4, 141.0, 132.3, 129.2, 128.8, 127.4, 127.2, 123.8, 122.8, 122.6, 122.5, 121.2, 118.3, 108.7, 105.8, 37.9, 13.8 ppm.
MS-EI (m/z) found 339 $[M]^+$, calcd 339.

Synthesis of 7-(3-(((*tert*-butyldimethylsilyl)oxy)methyl)thiophen-2-yl)-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.30o**)



Synthesised according to general procedure F. ((2-Bromothiophen-3-yl)methoxy)(tert-butyl)dimethylsilane (**4.38**) (300 mg, 0.97 mmol), 9-ethyl-7-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-9H-carbazole-3-carbaldehyde (**4.37**) (511 mg, 1.46 mmol), potassium carbonate (268 mg, 1.94 mmol), anhydrous MeCN (5 mL) and degassed H_2O (5 mL). Product was isolated as a beige solid, yield 70% (309 mg, 0.68 mmol).
^{1}H NMR (500 MHz, Chloroform-*d*) δ 10.10 (s, 1H, CHO), 8.60 (s, 1H, ar CH), 8.15 (d, *J* = 8.0 Hz, 1H, ar CH), 8.02 (d, *J* = 8.5 Hz, 1H, ar CH), 7.67 (s, 1H, ar CH), 7.50 – 7.44 (m, 2H, ar CH), 7.31 (d, *J* = 5.2 Hz, 1H, ar CH), 7.22 (d, *J* = 5.2 Hz, 1H, ar CH), 4.79 (s, 2H, CH_2), 4.41 (q, *J* = 7.2 Hz, 2H, CH_2), 1.48 (t, *J* = 7.2 Hz, 3H, CH_3), 0.96 (s, 9H, CH_3), 0.12 (s, 6H, CH_3) ppm.
^{13}C NMR (126 MHz, Chloroform-*d*) δ 191.7 (CHO), 144.0, 142.1, 140.8, 136.8, 132.2, 129.5, 128.7, 127.5, 124.7, 123.9, 122.6, 121.8, 120.9, 120.8, 109.8, 108.9, 58.9, 38.0, 25.6 (3C), 18.1, 13.8, -2.96 (2C) ppm.
HRMS-ESI (m/z) found 472.1748 $[M+Na]^+$, calcd 472.1737 for $[C_{26}H_{31}NOSSi+Na]^+$.

Synthesis of 1-(9-ethyl-7-(thiophen-3-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31c**)



Synthesised according to general procedure B. 9-Ethyl-7(thiophen-3-yl)-9*H*-carbazole-3-carbaldehyde (**4.30c**) (200 mg, 0.65 mmol), methylamine (491 µL, 2M solution in THF, 0.98 mmol), anhydrous THF (10 mL) and sodium triacetoxyborohydride (346 mg, 1.63 mmol). Product was isolated as an orange solid, yield 22 % (47 mg, 0.14 mmol).

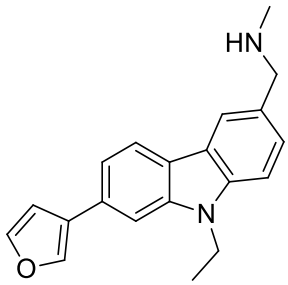
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.15 – 8.08 (m, 2H, ar CH), 7.59 (s, 1H, ar CH), 7.58 – 7.52 (m, 3H, ar CH), 7.52 – 7.47 (m, 1H, ar CH), 7.47 – 7.42 (m, 1H, ar CH), 7.41 – 7.36 (m, 1H, ar CH), 4.41 (q, *J* =7.0 Hz, 2H, CH_2), 3.83 (s, 2H, CH_2), 2.35 (s, 3H, CH_3), 1.47 (t, *J* = 7.0 Hz, 3H, CH_3) pp,.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 143.4, 140.7, 139.9, 133.6, 127.2, 126.8, 126.0, 122.8, 122.0, 121.1, 120.7, 120.0, 117.9, 108.2, 106.2, 61.9, 41.8, 37.5, 13.8 ppm.

HRMS-ESI (*m/z*) found 290.0997 $[M]^+$, calcd 290.0998 $[C_{19}H_{16}NS]^+$ (loss of NHMe).

LC-MS purity > 99% (UV), ret. time = 16.60 min.

Synthesis of 1-(9-ethyl-7-(furan-3-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31d**)



Synthesised according to general procedure B. 9-Ethyl-7-(furan-3-yl)-9*H*-carbazole-3-carbaldehyde (**4.30d**) (200 mg, 0.69 mmol), methylamine (518 µL, 2M solution in THF, 1.03 mmol), anhydrous THF (10 mL) and sodium triacteoxyborohydride (366 mg, 1.72 mmol). Product was isolated as an orange solid, yield 61% (129 mg, 0.42 mmol).

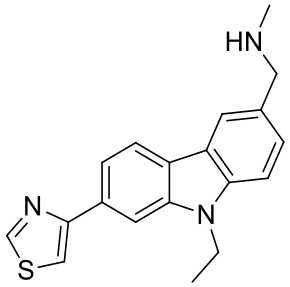
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.04 (d, *J* = 8.0 Hz, 1H, ar CH), 8.02 (s, 1H, ar CH), 7.82 (s, 1H, ar CH), 7.52 (s, 1H, ar CH), 7.44 (s, 1H, ar CH), 7.41 (d, *J* = 8.3 Hz, 1H, ar CH), 7.36 (d, *J* = 8.0 Hz, 1H, ar CH), 7.33 (d, *J* = 8.3 Hz, 1H, ar CH), 6.82 (s, 1H, ar CH), 4.34 (q, *J* = 7.2 Hz, 2H, CH_2), 3.94 (s, 2H, CH_2), 2.51 (s, 3H, CH_3), 1.42 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 143.5, 140.7, 139.7, 138.4, 130.1, 127.4, 126.3, 122.9, 121.9, 120.7, 120.3, 117.3, 109.9, 109.2, 108.3, 105.5, 55.6, 37.5, 34.9, 13.7 ppm.

HRMS-ESI (*m/z*) found 274.1222 $[M]^+$, calcd. 274.1226 for $[C_{19}H_{16}NO]^+$ (loss of NHMe).

LC-MS purity = 90 % (UV), ret. time = 12.68 min.

Synthesis of 1-(9-ethyl-7-(thiazol-4-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31e**)



Synthesised according to general procedure B. 9-Ethyl-7-(thiazol-4-yl)-9H-carbazole-3-carbaldehyde (**4.30e**) (342 mg, 1.11 mmol), methylamine (836 µL, 1.67 mmol), anhydrous THF (15 mL) and sodium triacetoxyborohydride (591 mg, 2.79 mmol). Product was isolated as a beige solid, yield 48% (172 mg, 0.53 mmol).

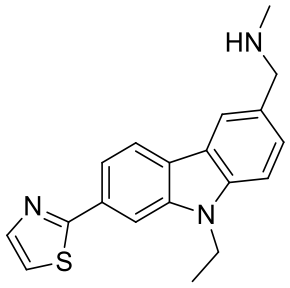
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.90 (d, *J* = 1.8 Hz, 1H, ar CH), 8.07 – 8.03 (m, 1H, ar CH), 8.03 – 7.99 (m, 2H, ar CH), 7.68 – 7.62 (m, 1H, ar CH), 7.55 (d, *J* = 1.7 Hz, 1H, ar CH), 7.42 – 7.36 (m, 1H, ar CH), 7.30 – 7.23 (m, 1H, ar CH), 4.28 (q, *J* = 7.2 Hz, 2H, CH_2), 3.94 (s, 2H, CH_2), 2.47 (s, 3H, CH_3), 1.35 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 157.1, 152.6, 140.6, 140.2, 131.9, 126.8, 126.6, 122.8, 122.6, 121.1, 120.7, 117.5, 112.2, 108.5, 106.7, 54.4, 37.5, 33.6, 13.8 ppm.

HRMS-ESI (*m/z*) found 291.0948 $[M]^+$, calcd. 291.0950 for $[C_{18}H_{15}N_2S]^+$ (loss of NHMe).

LC-MS purity > 99% (UV), ret. time = 11.35 min.

Synthesis of 1-(9-ethyl-7-(thiazol-2-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31f**)



Synthesised according to general procedure B. 9-Ethyl-7-(thiazol-2-yl)-9H-carbazole-3-carbaldehyde (**4.30f**) (130 mg, 0.42 mmol), methylamine (315 µL, 2M solution in THF, 0.63 mmol), anhydrous THF (7 mL) and sodium triacetoxyborohydride (224 mg, 1.06 mmol). Product was isolated as a beige solid, yield 10% (13 mg, 0.04 mmol).

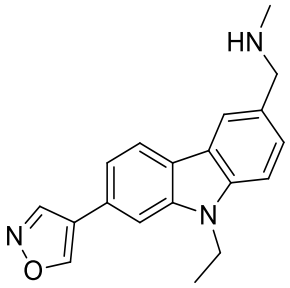
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.15 (s, 1H, ar CH), 8.03 (d, *J* = 8.0 Hz, 1H, ar CH), 7.93 (s, 1H, ar CH), 7.87 (d, *J* = 3.2 Hz, 1H, ar CH), 7.66 (d, *J* = 8.0 Hz, 1H, ar CH), 7.56 (d, *J* = 7.6 Hz, 1H, ar CH), 7.31 (d, *J* = 3.2 Hz, 1H, ar CH), 7.26 (d, *J* = 8.0 Hz, 1H, ar CH), 4.21 (q, *J* = 6.9 Hz, 2H, CH_2), 4.07 (s, 2H, CH_2), 2.53 (s, 3H, CH_3) and 1.29 (t, *J* = 6.9 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 169.1, 143.5, 141.0, 140.3, 131.4, 128.1, 123.8, 122.9, 122.7, 121.0, 120.6, 118.6, 118.4, 109.1, 106.2, 52.6, 37.6, 31.2, 13.7 ppm.

HRMS-ESI (*m/z*) found 291.0941 $[M]^+$, calcd 291.0950 for $[C_{18}H_{15}N_2S]^+$ (loss of NHMe).

LC-MS purity > 99 % (UV), ret. time = 11.93 min.

Synthesis of 1-(9-ethyl-7-(isoxazol-4-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanaime (**4.31g**)



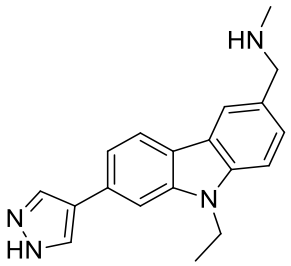
Synthesised according to general procedure B. 9-Ethyl-7-(isoxazol-4-yl)-9*H*-carbazole-3-carbaldehyde (**4.30h**) (200 mg, 0.68 mmol), methylamine (516 µL, 2M solution in

THF, 1.03 mmol), anhydrous THF (10 mL) and sodium triacetoxyborohydride (365 mg, 1.72 mmol). Product was isolated as a brown oil, yield 18% (38 mg, 0.12 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.68 (s, 1H, ar CH), 8.58 (s, 1H, ar CH), 8.04 (s, 1H, ar CH), 7.95 (d, *J* = 8.0 Hz, 1H, ar CH), 7.43 (d, *J* = 8.3 Hz, 1H, ar CH), 7.28 (d, *J* = 7.9 Hz, 2H, ar CH), 7.19 (d, *J* = 7.9 Hz, 1H, ar CH), 4.23 (q, *J* = 7.1 Hz, 2H, CH_2), 4.07 (s, 2H, CH_2), 2.54 (s, 3H, CH_3), 1.34 (t, *J* = 7.1 Hz, 2H, CH_3) ppm.

HRMS-ESI (*m/z*) found 275.1181 $[M]^+$, calcd 275.1179 for $[C_{18}H_{15}N_2O]^+$ (loss of NHMe).

Synthesis of 1-(9-Ethyl-7-(1*H*-pyrazol-4-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanaime (**4.31h**)



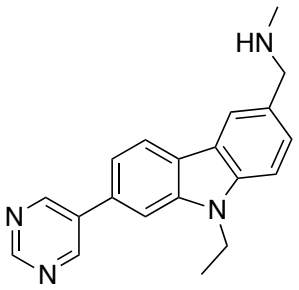
Synthesised according to general procedure B. 9-Ethyl-7-(1*H*-pyrazol-4-yl)-9*H*-carbazole-3-carbaldehyde (**4.30h**) (45 mg, 0.15 mmol), methylamine (116 μL, 2M solution in THF, 0.23 mmol), anhydrous THF (2 mL) and sodium triacetoxyborohydride (82 mg, 0.23 mmol). Product was isolated as a white solid, yield 29% (14 mg, 0.045 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.16 – 8.10 (m, 4H), 8.08 (s, 1H), 7.70 (d, *J* = 8.1 Hz, 1H), 7.40 (d, *J* = 8.3 Hz, 1H), 6.85 (d, *J* = 2.8 Hz, 1H), 4.44 (q, *J* = 7.1 Hz, 2H), 3.84 (s, 2H), 2.35 (s, 3H), 1.47 (t, *J* = 7.1 Hz, 3H) ppm.

HRMS-ESI *(m/z)* found 274.1340 $[M]^+$, calcd 274.1339 for $[C_{18}H_{16}N_3]^+$ (loss of NHMe).

LC-MS purity = 96 % (UV), ret. time = 13.13 min.

Synthesis of 1-(9-ethyl-7-(pyrimidin-5-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31i**)



Synthesised according to general procedure A. 9-Ethyl-7-(pyrimidin-5-yl)-9*H*-carbazole-3-carbaldehyde (**4.30i**) (21 mg, 0.06 mmol) methylamine hydrochloride (9 mg, 0.13 mmol), triethylamine (14 µL, 0.10 mmol), titanium (IV) isopropoxide (41 µL, 0.13 mmol), DCM (1 mL), EtOH (1 mL) and sodium borohydride (5 mg, 0.13 mmol). Product was isolated as an orange oil in a quantitative yield (22 mg, 0.06 mmol).

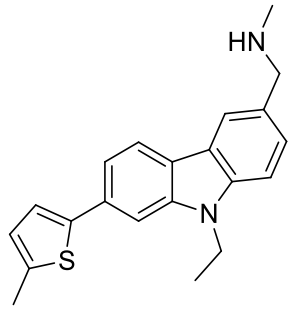
^{1}H NMR (500 MHz, DMSO-d_6) δ 9.28 (s, 2H, ar CH), 9.17 (s, 1H, ar CH), 8.24 (d, J = 7.8 Hz, 1H, ar CH), 8.17 (s, 1H, ar CH), 8.07 (s, 1H, ar CH), 7.64 – 7.54 (m, 2H, ar CH), 7.49 (d, J = 9.5 Hz, 1H, ar CH), 4.51 (q, J = 7.0 Hz, 2H, CH_2), 3.93 (s, 2H, CH_2), 2.39 (s, 3H, CH_3), 1.33 (t, J = 6.8 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 157.1, 155.1 (2C), 140.6, 140.0, 135.24, 131.6, 127.7 127.2, 123.3, 121.4, 121.3, 120.8, 117.8, 108.6, 106.7, 55.6, 40.6, 37.7, 13.8 ppm.

HRMS-ESI (m/z) found 286.1334 $[M]^+$, calcd $[C_{19}H_{16}N_3]^+$ (loss of NHMe).

LC-MS purity = 92 % (UV), ret. time = 10.86 min.

Synthesis of 1-(9-Ethyl-7-(5-methylthiophen-2-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31j**)



Synthesised according to general procedure B. 9-Ethyl-7-(5-methylthiophen-2-yl)-9H-carbazole-3-carbaldehyde (**4.30j**) (73 mg, 0.23 mmol), methylamine (171 µL, 2M solution in THF, 0.34 mmol), anhydrous THF (2 mL) and sodium triacetoxyborohydride

(122 mg, 0.57 mmol). Product was isolated as a pale yellow solid, yield 43% (33 mg, 0.09 mmol).

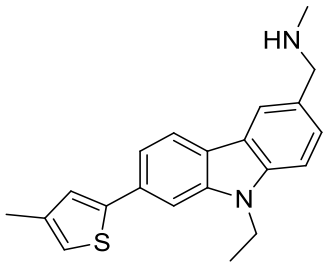
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.03 – 7.99 (m, 2H, ar CH), 7.49 (s, 1H, ar CH), 7.44 (d, *J* = 8.1 Hz, 1H, ar CH), 7.41 (d, *J* = 8.3 Hz, 1H, ar CH), 7.30 (d, *J* = 8.2 Hz, 1H, ar CH), 7.20 (d, *J* = 3.4 Hz, 1H, ar CH), 6.78 (d, *J* = 2.3 Hz, 1H, ar CH), 4.31 (q, *J* = 7.2 Hz, 2H, CH_2), 3.94 (s, 2H, CH_2), 2.55 (s, 3H, CH_3), 2.50 (s, 2H, CH_2), 1.40 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 142.3, 140.9, 140.6, 139.3, 132.6, 128.1, 127.2, 124.1, 122.8, 122.6, 122.4, 121.4, 121.2, 117.4, 109.7, 105.8, 52.1, 37.5, 32.2, 15.5, 14.1 ppm.

HRMS-ESI (*m/z*) found 304.1154 $[M]^+$, calcd 304.1154 for $[C_{20}H_{18}NS]^+$ (loss of NHMe).

LC-MS purity = 92% (UV), ret. time = 13.69 min.

Synthesis of 1-(9-Ethyl-7-(4-methylthiophen-2-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31k**)



Synthesised according to general procedure B. 9-Ethyl-7-(4-methylthiophene-2-yl)-9*H*-carbazole-3-carbaldehyde (**4.30k**) (285 mg, 0.89 mmol), methylamine (668 μL, 2M solution in THF, 1.33 mmol), anhydrous THF (5 mL), sodium triacetoxyborohydride (472 mg, 2.23 mmol). Product was isolated as a pale yellow solid, yield 28% (85 mg, 0.25 mmol).

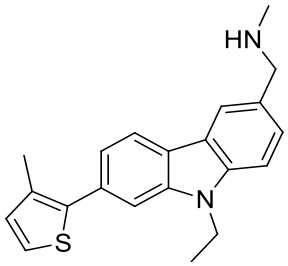
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.05 – 8.00 (m, 2H, ar CH), 7.54 (s, 1H, ar CH), 7.49 – 7.45 (m, 1H, ar CH), 7.42 (d, *J* = 8.3 Hz, 1H, ar CH), 7.32 (d, *J* = 8.3 Hz, 1H, ar CH), 7.23 (s, 1H, ar CH), 6.89 (s, 1H, ar CH), 4.34 (q, *J* = 7.2 Hz, 2H, CH_2), 3.94 (s, 2H, CH_2), 2.51 (s, 2H, CH_2), 2.34 (s, 3H, CH_3), 1.42 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, $cdcl_3$) δ 145.24, 140.65, 140.05, 138.62, 132.43, 126.60, 125.45, 122.90, 122.17, 120.73, 120.68, 119.92, 117.50, 108.44, 105.50, 55.14, 37.56, 34.39, 15.86, 13.77 ppm. Missing one quaternary carbon.

HRMS-ESI (*m/z*) found 304.1157 $[M]^+$, calcd 304.1154 for $[C_{20}H_{18}NS]^+$ (loss of NHMe).

LC-MS purity > 99% (UV), ret. time = 13.69 min.

Synthesis of 1-(9-Ethyl-7-(3-methylthiophen-2-yl)-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31l**)



9-Ethyl-7-(3-methylthiophen-2-yl)-9H-carbazole-3-carbaldehyde (**4.30l**) (120 mg, 0.37 mmol), methylamine (282 µL, 2M solution in THF, 0.56 mmol), anhydrous THF (4 mL) and sodium triacetoxyborohydride (199 mg, 0.93 mmol). Product was isolated as an orange solid, yield 39% (48 mg, 0.14 mmol).

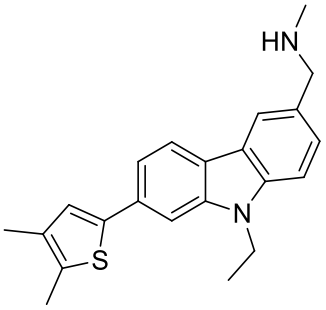
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.09 (d, *J* = 8.0 Hz, 1H, ar CH), 8.07 (s, 1H, ar CH), 7.44 (d, *J* = 7.3 Hz, 2H, ar CH), 7.34 (d, *J* = 8.2 Hz, 2H, ar CH), 7.25 (d, *J* = 5.1 Hz, 1H, ar CH), 6.98 (d, *J* = 5.1 Hz, 1H, ar CH), 4.33 (q, *J* = 7.2 Hz, 2H, CH_2), 3.98 (s, 2H, CH_2), 2.50 (s, 2H, CH_2), 2.41 (s, 3H, CH_3), 1.41 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 140.6, 140.4, 138.2, 133.4, 132.4, 131.9, 128.3, 124.6, 122.8, 122.2, 121.4, 120.9, 120.5, 117.7, 109.7, 109.6, 52.0, 37.5, 32.0, 15.3, 14.1 ppm.

MS-EI (m/z) found 335 $[M+H]^+$, calcd 335.

LC-MS purity 96% (UV), ret. time = 13.22 min.

Synthesis of 1-(7-(4,5-dimethylthiophen-2-yl)-9-ethyl-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31m**)



Synthesised according to general procedure B. 7-(4,5-Dimethylthiophen-2-yl)-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.30m**) (118 mg, 0.35 mmol), methylamine (265 µL, 2M solution in THF, 0.53 mmol), anhydrous THF (2 mL) and sodium triacetoxyborohydride

(187 mg, 0.88 mmol). Product was isolated as a pale brown solid, yield 46 % (59 mg, 0.17 mmol).

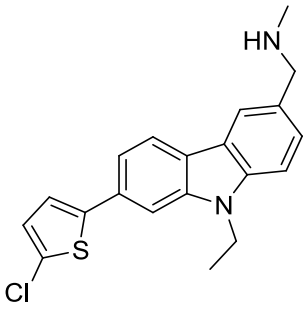
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.08 (s, 1H, ar CH), 8.04 (d, *J* = 7.7 Hz, 1H, ar CH), 7.58 (d, *J* = 8.0 Hz, 2H, ar CH), 7.52 (s, 1H, ar CH), 7.45 (d, *J* = 8.0 Hz, 1H, ar CH), 7.39 (d, *J* = 8.0 Hz, 1H, ar CH), 7.14 (s, 1H, ar CH), 4.37 (q, *J* = 6.3 Hz, 2H, CH_2), 4.05 (s, 2H, CH_2), 2.48 – 2.36 (m, 6H, CH_3), 2.21 (s, 3H, CH_3), 1.46 (t, *J* = 6.7 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 140.9, 140.8, 139.4, 134.5, 132.7, 132.4, 129.2, 127.0, 124.0, 122.5, 121.3, 120.7, 117.7, 117.2, 109.7, 105.5, 59.2, 38.1, 37.5, 14.2, 13.9, 13.3 ppm.

HRMS-ESI (*m/z*) found 318.1312 $[M]^+$, calcd 318.1311 for $[C_{21}H_{20}NS]^+$ (loss of NHMe).

LC-MS purity = 95% (UV), ret. time = 18.57 min.

Synthesis of 1-(7-(5-chlorothiophen-2-yl)-9-ethyl-9*H*-carbazol-3-yl)-*N*-methylmethanamine (**4.31n**)



7-(5-Chlorothiophen-2-yl)-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.30n**) (96 mg, 0.28 mmol), methylamine (211 μL, 2M solution in THF, 0.42 mmol), anhydrous THF (2 mL) and sodium triacetoxyborohydride (149 mg 0.70 mmol). Product was isolated as a bright yellow solid, yield 14% (14 mg, 0.039 mmol).

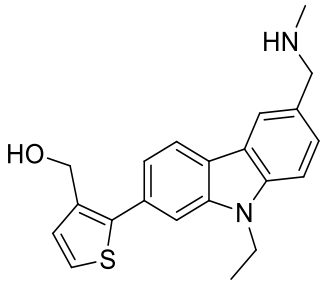
^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.04 – 7.98 (m, 3H, ar CH), 7.43 (d, *J* = 8.3 Hz, 2H, ar CH), 7.41 (s, 1H, ar CH), 7.36 (d, *J* = 8.2 Hz, 1H, ar CH), 7.30 (d, *J* = 8.3 Hz, 1H, ar CH), 7.14 (d, *J* = 3.8 Hz, 1H, ar CH), 6.93 (d, *J* = 3.8 Hz, 2H, ar CH), 4.29 (q, *J* = 7.2 Hz, 2H, CH_2), 3.95 (s, 2H, CH_2), 2.50 (s, 3H, CH_3), 1.38 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

^{13}C NMR (126 MHz, DMSO-d_6) δ 143.8, 140.8, 131.3, 128.6, 128.4, 127.4, 124.0, 122.9, 122.8, 122.2, 122.1, 121.4, 117.7, 117.4, 109.8, 106.3, 52.1, 37.5, 32.2, 14.1 ppm.

MS-EI (m/z) found 354 $[M]^+$, calcd 354.

LC-MS purity = 94% (UV), ret. time = 13.73 min.

Synthesis of (2-(9-ethyl-6-((methylamino)methyl)-9*H*-carbazol-2-yl)thiophen-3-yl)methanol (**4.31o**)



Synthesised according to general procedure B. 7-(3-(((*tert*-Butyldimethylsilyl)oxy)methyl)thiophen-2-yl)-9-ethyl-9*H*-carbazole-3-carbaldehyde (**4.30o**) (270 mg, 0.6 mmol), methylamine (450 μL, 2M solution in THF, 0.9 mmol), anhydrous THF (10 mL) and sodium triacetoxyborohydride (317 mg, 1.5 mmol). Crude product was purified by chromatography on silica gel using DCM/MeOH 9/1 as an eluent to yield 1-(7-(3-(((tert-butyldimethylsilyl)oxy)methyl)thiophen-2-yl)-9-ethyl-9H-carbazol-3-yl)-N-methylmethanamine as a brown oil, yield 26% (74 mg, 0.15 mmol).

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.11 – 8.08 (m, 1H, ar CH), 8.08 (s, 1H, ar CH), 7.57 (s, 1H, ar CH), 7.47 – 7.40 (m, 1H, ar CH), 7.38 – 7.34 (m, 3H, ar CH), 7.29 (s, 1H, ar CH), 7.29 – 7.25 (m, 1H, ar CH), 7.22 – 7.17 (m, 1H, ar CH), 4.78 (s, 2H, CH_2), 4.36 (q, *J* = 7.2 Hz, 2H, CH_2), 3.97 (s, 2H, CH_2), 2.52 (s, 3H, CH_3), 1.43 (t, *J* = 7.2 Hz, 3H, CH_3), 0.94 (s, 9H, CH_3), 0.08 (s, 6H, CH_3) ppm.

^{13}C NMR (126 MHz, Chloroform-*d*) δ 141.2, 140.3, 139.9, 137.2, 131.5, 129.6, 128.8, 126.7, 123.7, 122.7, 122.2, 120.6, 120.5, 120.3, 109.1, 108.4, 59.6, 55.4, 37.6, 34.7, 25.9 (3C), 18.3, 13.7, -5.1 (2C) ppm.

HRMS-ESI (*m/z*) found 434.1963 $[M]^+$, calcd. 434.1968 for $[C_{26}H_{32}NOSSi]^+$ (loss of NHMe).

1-(7-(3-(((tert-butyldimethylsilyl)oxy)methyl)thiophen-2-yl)-9-ethyl-9H-carbazol-3-yl)-N-methylmethanamine (70 mg, 0.15 mmol) was dissolved in anhydrous THF (1 mL) and cooled to 0 °C. TBAF (300 μL, 1M solution in THF, 0.30 mmol) was added dropwise and the solution was allowed to warm to room temperature. H_2O (10 mL) and EtOAc (10 mL) was added and product was extracted with EtOAc (3 X 10 mL). The organic extracts were washed with H_2O, dried over anhydrous $MgSO_4$, filtered and the filtrate was

concentrated under a reduced pressure to yield product that was purified by trituration with DCM. Product was isolated as a pale yellow solid, yield 19% (10 mg, 0.028 mmol). Successful removal of the silyl group was confirmed by ^{1}H NMR and mass spectrometry.

^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.20 (s, 1H, ar CH), 8.09 (d, *J* = 8.0 Hz, 1H, ar CH), 7.62 (d, *J* = 8.6 Hz, 1H, ar CH), 7.49 (s, 1H, ar CH), 7.36 – 7.29 (m, 3H, ar CH), 7.23 (d, *J* = 5.2 Hz, 1H, ar CH), 4.70 (s, 2H, CH_2), 4.28 – 4.18 (m, 2H, CH_2), 4.13 (s, 2H, CH_2), 2.53 (s, 3H, CH_3), 1.33 (t, *J* = 7.2 Hz, 3H, CH_3) ppm.

HRMS-ESI (*m/z*) found 320.1097 $[M]^+$, calcd. 320.1104 for $[C_{20}H_{18}NOS]^+$ (loss of NHMe).

LC-MS purity = 94% (UV), ret. time = 12.53 min.

Assay Conditions

All DSF, ITC and cell viability measurements were performed by Dr. Matthias Bauer at the MRC Laboratory of Molecular Biology according to these conditions.

DSF

The effect of compounds on the melting temperature of T-p53C-Y220C was monitored using SYPRO Orange (Invitrogen) as the fluorescent probe, which quantitatively binds to the hydrophobic protein patches exposed upon thermal denaturation. Real-time melt analysis was performed using a Corbett Rotor-Gene 6000 real-time qPCR thermocycler. Excitation and emission filters were set to 460 and 510 nm, respectively. Heating from 28 to 60 °C, a constant heating rate of 270 K/h was applied. The protein (final concentration of 10 μM) was briefly mixed with SYPRO orange (10×) in buffer (25 mM KPi pH 7.2, 150 mM NaCl, 1 mM TCEP), and compound (5 mM) dissolved in DMSO was added to give a final compound concentration of 250 μM in 5% (v/v) DMSO. The melting temperature (T_m) of the protein (10 μM) in presence of compounds was determined from the inflection point of the melting curve. Melting temperatures were compared with control samples without compound (yielding ΔT_mDSF). All samples were measured in triplicate.

ITC

ITC experiments were conducted using a MicroCal (Amherst) iTC200 calorimeter. Protein samples used in the cell unit were prepared to a final concentration of 50−200 μM

in 25 mM KPi, pH 7.2, 150 mM NaCl, 1 mM TCEP in 5% (v/v) DMSO. Compounds for use in the syringe unit were dissolved in the same buffer at 5% (v/v) DMSO. Measurements were performed at 20 °C using injection steps of 2 µL at a rate of 0.5 µL/s (initial injection: 0.5 µL) and 120 s spacing. Data analysis was performed using MicroCal Origin software.

NMR Spectroscopy

$^{1}H/^{15}N$-HSQC spectra of uniformly ^{15}N-labelled T-p53-Y220C (75 µM) with and without compounds were acquired at 20 °C on a Bruker Avance-800 spectrometer using a 5-mm inverse cryogenic probe. Samples were prepared by adding dilutions of compound from stock solutions in DMSO-d_6 to a final concentration of 5% (v/v) DMSO-d_6 in buffer. All HSQC spectra were acquired with 8 transients per t_1 data point, 1024 data points in t_2, and 64 complex data points in t_1, with spectral widths of 11.0 kHz for ^{1}H and 2.7 kHz for ^{15}N, and a recycle delay of 800 ms. After zero filling, forward complex linear prediction in f_1 and Fourier transformation, the digital resolytion was 0.01 ppm/point for ^{1}H and 0.13 ppm/point for ^{15}N. Chemicals shifts were considered significant if the average weighted $^{1}H/^{15}N$ chemical shift difference $\Delta\delta(^{1}H/^{15}N) = [(\Delta\delta(^{1}H))^2 + (\Delta\delta(^{15}N)/5)^2]^{1/2}$ was greater than 0.04 ppm. To determine dissociation constants, at least five $^{15}N/^{1}H$ HSQC spectra at different compound concentrations were measured. Spectra analysis was performed using Sparky 3.114 and Bruker Topspin 2.0 software. To derive K_D values, a quadratic saturation binding equation was fitted to the concentration-dependent chemical shift changes of the relevant shifting peaks:

$$\delta_{obs} = \delta_{max} = \frac{[L_0] + [P_0] + K_D - \sqrt{([L_0] + [P_0] + K_D)^2 - 4 \cdot [P_0] \cdot [L_0]}}{2 \cdot [P_0]}$$

Cell Viability

Cell viability was measured using the CellTiter-Glo 2.0 assay kit (Promega, USA) according to the kit instructions. Cells were seeded in 96 well plates at 5000 cells per well and incubated overnight. Samples were prepared in medium with a twice as high compound and DMSO concentration then added to an equivalent volume of growth medium, yielding a final DMSO concentration of 0.5%. After incubating of cells for 72 hours and equilibration of the 96-well plate to room temperature for 30 minutes,

CellTiter-Glo reagent was added to each well. Luminescence was recorded with a Centro XS³ LB 960 microplate luminometer (Berthold Technologies, Germany). Experiments were performed in quadruplicate.

Chapter 7 : References

1. A. S. Ahmad, N. Ormiston-Smith, and P. D. Sasieni, *Br. J. Cancer*, 2015, **112**, 943–947.
2. D. Hanahan and R. A. Weinberg, *Cell*, 2011, **144**, 646–74.
3. N. Rahman and M. R. Stratton, *Annu. Rev. Genet.*, 1998, **32**, 95–121.
4. P. C. Fong, D. S. Boss, T. A. Yap, A. Tutt, P. Wu, M. Mergui-Roelvink, P. Mortimer, H. Swaisland, A. Lau, M. J. O'Connor, A. Ashworth, J. Carmichael, S. B. Kaye, J. H. M. Schellens, and J. S. de Bono, *N. Engl. J. Med.*, 2009, **361**, 123–134.
5. D. P. Lane and L. V Crawford, *Nature*, 1979, **278**, 261–3.
6. K. H. Vousden and C. Prives, *Cell*, 2009, **137**, 413–431.
7. D. P. Lane, *Nature*, 1992, **358**, 15–16.
8. A. C. Joerger and A. R. Fersht, *Cold Spring Harb. Perspect. Biol.*, 2010, **2**, a000919–a000919.
9. Y. Cho, S. Gorina, P. D. Jeffrey, and N. P. Pavletich, *Science*, 1994, **265**, 346–55.
10. J. M. P. Canadillas, H. Tidow, S. M. V. Freund, T. J. Rutherford, H. C. Ang, and A. R. Fersht, *Proc. Natl. Acad. Sci.*, 2006, **103**, 2109–2114.
11. Y. Wang, A. Rosengarth, and H. Luecke, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2007, **63**, 276–281.
12. A. N. Bullock, J. Henckel, B. S. DeDecker, C. M. Johnson, P. V Nikolova, M. R. Proctor, D. P. Lane, and A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 14338–42.
13. A. Friedler, D. B. Veprintsev, L. O. Hansson, and A. R. Fersht, *J. Biol. Chem.*, 2003, **278**, 24108–24112.
14. H. C. Ang, A. C. Joerger, S. Mayer, and A. R. Fersht, *J. Biol. Chem.*, 2006, **281**, 21934–21941.
15. K. H. Khoo, A. Andreeva, and A. R. Fersht, *J. Mol. Biol.*, 2009, **393**, 161–175.
16. K. H. Khoo, A. C. Joerger, S. M. V. Freund, and A. R. Fersht, *Protein Eng. Des. Sel.*, 2009, **22**, 421–430.
17. A. R. Fersht, M. G. Mateu, and M. M. Sánchez Del Pino, *Nat. Struct. Biol.*, 1999, **6**, 191–198.
18. C. D. Nicholls, K. G. McLure, M. A. Shields, and P. W. K. Lee, *J. Biol. Chem.*, 2002, **277**, 12937–12945.

19. M. G. Mateu and A. R. Fersht, *EMBO J.*, 1998, **17**, 2748–58.
20. P. Mora, R. J. Carbajo, A. Pineda-Lucena, M. M. Sánchez del Pino, and E. Pérez-Payá, *Proteins Struct. Funct. Bioinforma.*, 2007, **71**, 1670–1685.
21. J. M. Stommel, N. D. Marchenko, G. S. Jimenez, U. M. Moll, T. J. Hope, and G. M. Wahl, *EMBO J.*, 1999, **18**, 1660–72.
22. R. S.-Y. Foo, Y.-J. Nam, M. J. Ostreicher, M. D. Metzl, R. S. Whelan, C.-F. Peng, A. W. Ashton, W. Fu, K. Mani, S.-F. Chin, E. Provenzano, I. Ellis, N. Figg, S. Pinder, M. R. Bennett, C. Caldas, and R. N. Kitsis, *Proc. Natl. Acad. Sci.*, 2007, **104**, 20826–20831.
23. S. Rajagopalan, A. M. Jaulent, M. Wells, D. B. Veprintsev, and A. R. Fersht, *Nucleic Acids Res.*, 2008, **36**, 5983–5991.
24. H. Lee, K. H. Mok, R. Muhandiram, K.-H. Park, J.-E. Suk, D.-H. Kim, J. Chang, Y. C. Sung, K. Y. Choi, and K.-H. Han, *J. Biol. Chem.*, 2000, **275**, 29426–29432.
25. R. Rosal, M. R. Pincus, P. W. Brandt-Rauf, R. L. Fine, A. Josef Michl, and H. Wang, *Biochemistry*, 2004, **43**, 1854–1861.
26. M. Wells, H. Tidow, T. J. Rutherford, P. Markwick, M. R. Jensen, E. Mylonas, D. I. Svergun, M. Blackledge, and A. R. Fersht, *Proc. Natl. Acad. Sci.*, 2008, **105**, 5762–5767.
27. P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine, and N. P. Pavletich, *Science*, 1996, **274**, 948–53.
28. G. Popowicz, A. Czarna, and T. Holak, *Cell Cycle*, 2008, **7**, 2441–2443.
29. E. Bochkareva, L. Kaustov, A. Ayed, G.-S. Yi, Y. Lu, A. Pineda-Lucena, J. C. C. Liao, A. L. Okorokov, J. Milner, C. H. Arrowsmith, and A. Bochkarev, *Proc. Natl. Acad. Sci.*, 2005, **102**, 15412–15417.
30. O. Schon, A. Friedler, M. Bycroft, S. M. V Freund, and A. R. Fersht, *J. Mol. Biol.*, 2002, **323**, 491–501.
31. C. J. Brown, D. Srinivasan, L. H. Jun, D. Coomber, C. S. Verma, and D. P. Lane, *Cell Cycle*, 2008, **7**, 608–610.
32. K. Sakaguchi, S. Saito, Y. Higashimoto, S. Roy, C. W. Anderson, and E. Appella, *J. Biol. Chem.*, 2000, **275**, 9278–83.
33. J. C. Ferreon, C. W. Lee, M. Arai, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, *Proc. Natl. Acad. Sci.*, 2009, **106**, 6591–6596.
34. L. M. M. Jenkins, H. Yamaguchi, R. Hayashi, S. Cherry, J. E. Tropea, M. Miller,

A. Wlodawer, E. Appella, and S. J. Mazur, *Biochemistry*, 2009, **48**, 1244–1255.

35. C. W. Lee, M. Arai, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, *Biochemistry*, 2009, **48**, 2115–2124.
36. D. P. Teufel, M. Bycroft, and A. R. Fersht, *Oncogene*, 2009, **28**, 2112–2118.
37. F. Toledo and G. M. Wahl, *Nat. Rev. Cancer*, 2006, **6**, 909–923.
38. J.-P. Kruse and W. Gu, *Cell*, 2009, **137**, 609–622.
39. M. R. Fernandez-Fernandez, T. J. Rutherford, and A. R. Fersht, *Protein Sci.*, 2008, **17**, 1663–1670.
40. J. L. Avalos, I. Celic, S. Muhammad, M. S. Cosgrove, J. D. Boeke, and C. Wolberger, *Mol. Cell*, 2002, **10**, 523–35.
41. A. J. Levine, *Cell*, 1997, **88**, 323–31.
42. K. S. Yee and K. H. Vousden, *Carcinogenesis*, 2005, **26**, 1317–1322.
43. E. Appella and C. W. Anderson, *Eur. J. Biochem.*, 2001, **268**, 2764–72.
44. C. J. Sherr, *Nat. Rev. Cancer*, 2006, **6**, 663–673.
45. W. S. El-Deiry, S. E. Kern, J. A. Pietenpol, K. W. Kinzler, and B. Vogelstein, *Nat. Genet.*, 1992, **1**, 45–49.
46. Y. Liu, J. P. Lagowski, G. E. Vanderbeek, and M. F. Kulesz-Martin, *Cancer Biol. Ther.*, 2004, **3**, 1102–8.
47. J. M. Espinosa, R. E. Verdun, and B. M. Emerson, *Mol. Cell*, 2003, **12**, 1015–27.
48. J. Chen, *Genes Cancer*, 2012, **3**, 274–82.
49. W. S. el-Deiry, T. Tokino, V. E. Velculescu, D. B. Levy, R. Parsons, J. M. Trent, D. Lin, W. E. Mercer, K. W. Kinzler, and B. Vogelstein, *Cell*, 1993, **75**, 817–25.
50. J. W. Harper, G. R. Adami, N. Wei, K. Keyomarsi, and S. J. Elledge, *Cell*, 1993, **75**, 805–16.
51. R. L. Weinberg, D. B. Veprintsev, M. Bycroft, and A. R. Fersht, *J. Mol. Biol.*, 2005, **348**, 589–596.
52. C. Deng, P. Zhang, J. W. Harper, S. J. Elledge, and P. Leder, *Cell*, 1995, **82**, 675–84.
53. J. Brugarolas, C. Chandrasekaran, J. I. Gordon, D. Beach, T. Jacks, and G. J. Hannon, *Nature*, 1995, **377**, 552–557.
54. J. Martín-Caballero, J. M. Flores, P. García-Palencia, and M. Serrano, *Cancer Res.*, 2001, **61**, 6234–8.
55. S. St. Clair, L. Giono, S. Varmeh-Ziaie, L. Resnick-Silverman, W. Liu, A. Padi, J.

Dastidar, A. DaCosta, M. Mattia, and J. J. Manfredi, *Mol. Cell*, 2004, **16**, 725–736.

56. J. P. Brown, W. Wei, and J. M. Sedivy, *Science*, 1997, **277**, 831–4.
57. C. M. Beausejour, A. Krtolica, F. Galimi, M. Narita, S. W. Lowe, P. Yaswen, and J. Campisi, *EMBO J.*, 2003, **22**, 4212–4222.
58. M. Narita, S. Nũnez, E. Heard, M. Narita, A. W. Lin, S. A. Hearn, D. L. Spector, G. J. Hannon, and S. W. Lowe, *Cell*, 2003, **113**, 703–16.
59. R. Bonasio, S. Tu, and D. Reinberg, *Science*, 2010, **330**, 612–616.
60. F. Rodier, D. P. Munoz, R. Teachenor, V. Chu, O. Le, D. Bhaumik, J.-P. Coppe, E. Campeau, C. M. Beausejour, S.-H. Kim, A. R. Davalos, and J. Campisi, *J. Cell Sci.*, 2011, **124**, 68–81.
61. B. Huang, D. Deo, M. Xia, and L. T. Vassilev, *Mol. Cancer Res.*, 2009, **7**, 1497–1509.
62. P. Sdek, H. Ying, D. L. F. Chang, W. Qiu, H. Zheng, R. Touitou, M. J. Allday, and Z.-X. Jim Xiao, *Mol. Cell*, 2005, **20**, 699–708.
63. W. Du, J. Wu, E. M. Walsh, Y. Zhang, C. Y. Chen, and Z.-X. J. Xiao, *J. Biol. Chem.*, 2009, **284**, 26315–26321.
64. Z. N. Demidenko, S. G. Zubova, E. I. Bukreeva, V. A. Pospelov, T. V. Pospelova, and M. V Blagosklonny, *Cell Cycle*, 2009, **8**, 1888–1895.
65. M. Müller, S. Wilder, D. Bannasch, D. Israeli, K. Lehlbach, M. Li-Weber, S. L. Friedman, P. R. Galle, W. Stremmel, M. Oren, and P. H. Krammer, *J. Exp. Med.*, 1998, **188**, 2033–45.
66. M. Bennett, *Science*, 1998, **282**, 290–293.
67. G. S. Wu, T. F. Burns, E. R. McDonald, W. Jiang, R. Meng, I. D. Krantz, G. Kao, D.-D. Gan, J.-Y. Zhou, R. Muschel, S. R. Hamilton, N. B. Spinner, S. Markowitz, G. Wu, and W. S. El-Deiry, *Nat. Genet.*, 1997, **17**, 141–143.
68. S. W. G. Tait and D. R. Green, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 621–632.
69. N. D. Marchenko, A. Zaika, and U. M. Moll, *J. Biol. Chem.*, 2000, **275**, 16202–12.
70. M. Mihara, S. Erster, A. Zaika, O. Petrenko, T. Chittenden, P. Pancoska, and U. M. Moll, *Mol. Cell*, 2003, **11**, 577–590.
71. A. M. Petros, A. Gunasekera, N. Xu, E. T. Olejniczak, and S. W. Fesik, *FEBS Lett.*, 2004, **559**, 171–174.
72. B. Sot, S. M. V. Freund, and A. R. Fersht, *J. Biol. Chem.*, 2007, **282**, 29193–29200.

73. J. I.-J. Leu, P. Dumont, M. Hafey, M. E. Murphy, and D. L. George, *Nat. Cell Biol.*, 2004, **6**, 443–450.
74. E. C. Pietsch, E. Perchiniak, A. A. Canutescu, G. Wang, R. L. Dunbrack, and M. E. Murphy, *J. Biol. Chem.*, 2008, **283**, 21294–304.
75. A. K. Dunker, I. Silman, V. N. Uversky, and J. L. Sussman, *Curr. Opin. Struct. Biol.*, 2008, **18**, 756–764.
76. A. R. Clarke, C. A. Purdie, D. J. Harrison, R. G. Morris, C. C. Bird, M. L. Hooper, and A. H. Wyllie, *Nature*, 1993, **362**, 849–852.
77. H. Symonds, L. Krall, L. Remington, M. Saenz-Robles, S. Lowe, T. Jacks, and T. Van Dyke, *Cell*, 1994, **78**, 703–11.
78. S. Rowan, R. L. Ludwig, Y. Haupt, S. Bates, X. Lu, M. Oren, and K. H. Vousden, *EMBO J.*, 1996, **15**, 827–38.
79. K. Schlereth, R. Beinoraviciute-Kellner, M. K. Zeitlinger, A. C. Bretz, M. Sauer, J. P. Charles, F. Vogiatzi, E. Leich, B. Samans, M. Eilers, C. Kisker, A. Rosenwald, and T. Stiewe, *Mol. Cell*, 2010, **38**, 356–368.
80. G. Liu, J. M. Parant, G. Lang, P. Chau, A. Chavez-Reyes, A. K. El-Naggar, A. Multani, S. Chang, and G. Lozano, *Nat. Genet.*, 2004, **36**, 63–68.
81. S. M. Post, A. Quintás-Cardama, T. Terzian, C. Smith, C. M. Eischen, and G. Lozano, *Oncogene*, 2010, **29**, 1260–1269.
82. O. Timofeev, K. Schlereth, M. Wanzel, A. Braun, B. Nieswandt, A. Pagenstecher, A. Rosenwald, H.-P. Elsässer, and T. Stiewe, *Cell Rep.*, 2013, **3**, 1512–1525.
83. C. A. Schmitt, J. S. Fridman, M. Yang, S. Lee, E. Baranov, R. M. Hoffman, and S. W. Lowe, *Cell*, 2002, **109**, 335–46.
84. S. P. Garrison, J. R. Jeffers, C. Yang, J. A. Nilsson, M. A. Hall, J. E. Rehg, W. Yue, J. Yu, L. Zhang, M. Onciu, J. T. Sample, J. L. Cleveland, and G. P. Zambetti, *Mol. Cell. Biol.*, 2008, **28**, 5391–5402.
85. W. Xue, L. Zender, C. Miething, R. A. Dickins, E. Hernando, V. Krizhanovsky, C. Cordon-Cardo, and S. W. Lowe, *Nature*, 2007, **445**, 656–60.
86. A. Ventura, D. G. Kirsch, M. E. McLaughlin, D. A. Tuveson, J. Grimm, L. Lintault, J. Newman, E. E. Reczek, R. Weissleder, and T. Jacks, *Nature*, 2007, **445**, 661–5.
87. Y. Wang, Y.-A. Suh, M. Y. Fuller, J. G. Jackson, S. Xiong, T. Terzian, A. Quintás-Cardama, J. A. Bankson, A. K. El-Naggar, and G. Lozano, *J. Clin. Invest.*, 2011,

121, 893–904.

88. T. Li, N. Kon, L. Jiang, M. Tan, T. Ludwig, Y. Zhao, R. Baer, and W. Gu, *Cell*, 2012, **149**, 1269–83.
89. L. J. Valente, D. H. D. Gray, E. M. Michalak, J. Pinon-Hofbauer, A. Egle, C. L. Scott, A. Janic, and A. Strasser, *Cell Rep.*, 2013, **3**, 1339–1345.
90. V. P. Sah, L. D. Attardi, G. J. Mulligan, B. O. Williams, R. T. Bronson, and T. Jacks, *Nat. Genet.*, 1995, **10**, 175–180.
91. L. A. Donehower, M. Harvey, B. L. Slagle, M. J. McArthur, C. A. Montgomery, J. S. Butel, and A. Bradley, *Nature*, 1992, **356**, 215–221.
92. S. Venkatachalam, Y. P. Shi, S. N. Jones, H. Vogel, A. Bradley, D. Pinkel, and L. A. Donehower, *EMBO J.*, 1998, **17**, 4657–67.
93. S. D. Tyner, S. Venkatachalam, J. Choi, S. Jones, N. Ghebraniousk, H. Igelmann, X. Lu, G. Soron, B. Cooper, C. Brayton, S. Hee ParkI, T. ThompsonI, G. Karsenty, A. Bradley, and L. A. Donehower, *Nature*, 2002, **415**.
94. A. Gentry and S. Venkatachalam, *Aging Cell*, 2005, **4**, 157–160.
95. B. Maier, W. Gluba, B. Bernier, T. Turner, K. Mohammad, T. Guise, A. Sutherland, M. Thorner, and H. Scrable, *Genes Dev.*, 2004, **18**, 306–19.
96. L. Guarente and C. Kenyon, *Nature*, 2000, **408**, 255–262.
97. I. García-Cao, M. García-Cao, J. Martín-Caballero, L. M. Criado, P. Klatt, J. M. Flores, J.-C. Weill, M. A. Blasco, and M. Serrano, *EMBO J.*, 2002, **21**, 6225–35.
98. S. M. Mendrysa, K. A. O'Leary, M. K. McElwee, J. Michalowski, R. N. Eisenman, D. A. Powell, and M. E. Perry, *Genes Dev.*, 2006, **20**, 16–21.
99. S. M. Mendrysa and M. E. Perry, *Cell Cycle*, 2006, **5**, 714–717.
100. H. S. Gannon, L. A. Donehower, S. Lyle, and S. N. Jones, *Dev. Biol.*, 2011, **353**, 1–9.
101. I. Shibagaki, H. Tanaka, Y. Shimada, T. Wagata, M. Ikenaga, M. Imamura, and K. Ishizaki, *Clin. Cancer Res.*, 1995, **1**, 769–73.
102. A. Forslund, Z. Zeng, L.-X. Qin, S. Rosenberg, M. Ndubuisi, H. Pincas, W. Gerald, D. A. Notterman, F. Barany, and P. B. Paty, *Mol. Cancer Res.*, 2008, **6**, 205–211.
103. L. Chin, M. Meyerson, and et al., *Nature*, 2008, **455**, 1061–1068.
104. S. Lam, K. Lodder, A. F. A. S. Teunisse, M. J. W. E. Rabelink, M. Schutte, and A. G. Jochemsen, *Oncogene*, 2010, **29**, 2415–2426.
105. S. Mejia-Guerrero, M. Quejada, N. Gokgoz, M. Gill, R. K. Parkes, J. S. Wunder,

and I. L. Andrulis, *Genes, Chromosom. Cancer*, 2010, **49**, NA-NA.

106. M. Ito, L. Barys, T. O'Reilly, S. Young, B. Gorbatcheva, J. Monahan, S. Zumstein-Mecker, P. F. Choong, I. Dickinson, P. Crowe, C. Hemmings, J. Desai, D. M. Thomas, and J. Lisztwan, *Clin. Cancer Res.*, 2011, **17**, 416–426.
107. D. M. Gilkes, Y. Pan, D. Coppola, T. Yeatman, G. W. Reuther, and J. Chen, *Mol. Cell. Biol.*, 2008, **28**, 1999–2010.
108. A. Gembarska, F. Luciani, C. Fedele, E. A. Russell, M. Dewaele, S. Villar, A. Zwolinska, S. Haupt, J. de Lange, D. Yip, J. Goydos, J. J. Haigh, Y. Haupt, L. Larue, A. Jochemsen, H. Shi, G. Moriceau, R. S. Lo, G. Ghanem, M. Shackleton, F. Bernal, and J.-C. Marine, *Nat. Med.*, 2012, **18**, 1239–1247.
109. K. I. Pishas, F. Al-Ejeh, I. Zinonos, R. Kumar, A. Evdokiou, M. P. Brown, D. F. Callen, and P. M. Neilsen, *Clin. Cancer Res.*, 2011, **17**, 494–504.
110. N. A. Laurie, S. L. Donovan, C.-S. Shih, J. Zhang, N. Mills, C. Fuller, A. Teunisse, S. Lam, Y. Ramos, A. Mohan, D. Johnson, M. Wilson, C. Rodriguez-Galindo, M. Quarto, S. Francoz, S. M. Mendrysa, R. Kiplin Guy, J.-C. Marine, A. G. Jochemsen, and M. A. Dyer, *Nature*, 2006, **444**, 61–66.
111. J. McEvoy, A. Ulyanov, R. Brennan, G. Wu, S. Pounds, J. Zhang, and M. A. Dyer, *PLoS One*, 2012, **7**, e42739.
112. J. Momand, G. P. Zambetti, D. C. Olson, D. George, and A. J. Levine, *Cell*, 1992, **69**, 1237–1245.
113. S. Fang, J. P. Jensen, R. L. Ludwig, K. H. Vousden, and A. M. Weissman, *J. Biol. Chem.*, 2000, **275**, 8945–51.
114. Y. Barak, E. Gottlieb, T. Juven-Gershon, and M. Oren, *Genes Dev.*, 1994, **8**, 1739–49.
115. A. Phillips, A. Teunisse, S. Lam, K. Lodder, M. Darley, M. Emaduddin, A. Wolf, J. Richter, J. de Lange, M. Verlaan-de Vries, K. Lenos, A. Böhnke, F. Bartel, J. P. Blaydes, and A. G. Jochemsen, *J. Biol. Chem.*, 2010, **285**, 29111–29127.
116. L. K. Linares, A. Hengstermann, A. Ciechanover, S. Muller, and M. Scheffner, *Proc. Natl. Acad. Sci.*, 2003, **100**, 12009–12014.
117. L. Huang, Z. Yan, X. Liao, Y. Li, J. Yang, Z.-G. Wang, Y. Zuo, H. Kawai, M. Shadfan, S. Ganapathy, and Z.-M. Yuan, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 12001–6.
118. V. Pant, S. Xiong, T. Iwakuma, A. Quintas-Cardama, and G. Lozano, *Proc. Natl.*

Acad. Sci., 2011, **108**, 11995–12000.

119. A. Plechanovová, E. G. Jaffray, M. H. Tatham, J. H. Naismith, and R. T. Hay, *Nature*, 2012, **489**, 115–120.
120. M. V Poyurovsky, C. Priest, A. Kentsis, K. L. B. Borden, Z.-Q. Pan, N. Pavletich, and C. Prives, *EMBO J.*, 2007, **26**, 90–101.
121. S. Uldrijan, W.-J. Pannekoek, and K. H. Vousden, *EMBO J.*, 2007, **26**, 102–112.
122. F. Pichiorri, S.-S. Suh, A. Rocci, L. De Luca, C. Taccioli, R. Santhanam, W. Zhou, D. M. Benson, C. Hofmainster, H. Alder, M. Garofalo, G. Di Leva, S. Volinia, H.-J. Lin, D. Perrotti, M. Kuehl, R. I. Aqeilan, A. Palumbo, and C. M. Croce, *Cancer Cell*, 2010, **18**, 367–381.
123. L. D. Mayo, J. J. Turchi, and S. J. Berberich, *Cancer Res.*, 1997, **57**, 5013–6.
124. V. Zuckerman, K. Lenos, G. M. Popowicz, I. Silberman, T. Grossman, J.-C. Marine, T. A. Holak, A. G. Jochemsen, and Y. Haupt, *J. Biol. Chem.*, 2009, **284**, 4031–4039.
125. D. W. Meek and T. R. Hupp, *Semin. Cancer Biol.*, 2010, **20**, 19–28.
126. R. Maya, M. Balass, S. T. Kim, D. Shkedy, J. F. Leal, O. Shifman, M. Moas, T. Buschmann, Z. Ronai, Y. Shiloh, M. B. Kastan, E. Katzir, and M. Oren, *Genes Dev.*, 2001, **15**, 1067–77.
127. L. Chen, D. M. Gilkes, Y. Pan, W. S. Lane, and J. Chen, *EMBO J.*, 2005, **24**, 3411–3422.
128. Y. Pereg, S. Lam, A. Teunisse, S. Biton, E. Meulmeester, L. Mittelman, G. Buscemi, K. Okamoto, Y. Taya, Y. Shiloh, and A. G. Jochemsen, *Mol. Cell. Biol.*, 2006, **26**, 6819–6831.
129. Q. Cheng, L. Chen, Z. Li, W. S. Lane, and J. Chen, *EMBO J.*, 2009, **28**, 3857–3867.
130. L. D. Mayo and D. B. Donner, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 11598–603.
131. V. Lopez-Pajares, M. M. Kim, and Z.-M. Yuan, *J. Biol. Chem.*, 2008, **283**, 13707–13.
132. Y. V. Wang, M. Leblanc, M. Wade, A. G. Jochemsen, and G. M. Wahl, *Cancer Cell*, 2009, **16**, 33–43.
133. H. S. Gannon, B. A. Woda, and S. N. Jones, *Cancer Cell*, 2012, **21**, 668–679.
134. K. Lundgren, R. Montes de Oca Luna, Y. B. McNeill, E. P. Emerick, B. Spencer,

C. R. Barfield, G. Lozano, M. P. Rosenberg, and C. A. Finlay, *Genes Dev.*, 1997, **11**, 714–25.

135. S. N. Jones, A. R. Hancock, H. Vogel, L. A. Donehower, and A. Bradley, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 15608–12.
136. S. De Clercq, A. Gembarska, G. Denecker, M. Maetens, M. Naessens, K. Haigh, J. J. Haigh, and J.-C. Marine, *Mol. Cell. Biol.*, 2010, **30**, 5394–5405.
137. S. Xiong, V. Pant, Y.-A. Suh, C. S. Van Pelt, Y. Wang, Y. A. Valentin-Vega, S. M. Post, and G. Lozano, *Cancer Res.*, 2010, **70**, 7148–7154.
138. D. Danovi, E. Meulmeester, D. Pasini, D. Migliorini, M. Capra, R. Frenk, P. de Graaf, S. Francoz, P. Gasparini, A. Gobbi, K. Helin, P. G. Pelicci, A. G. Jochemsen, and J.-C. Marine, *Mol. Cell. Biol.*, 2004, **24**, 5835–43.
139. K. Lenos, J. de Lange, A. F. A. S. Teunisse, K. Lodder, M. Verlaan-de Vries, E. Wiercinska, M. J. M. van der Burg, K. Szuhai, and A. G. Jochemsen, *Mol. Cancer*, 2011, **10**, 111.
140. D. E. Quelle, F. Zindy, R. A. Ashmun, and C. J. Sherr, *Cell*, 1995, **83**, 993–1000.
141. T. Kamijo, F. Zindy, M. F. Roussel, D. E. Quelle, J. R. Downing, R. A. Ashmun, G. Grosveld, and C. J. Sherr, *Cell*, 1997, **91**, 649–59.
142. R. Honda and H. Yasuda, *EMBO J.*, 1999, **18**, 22–27.
143. M. Ruas and G. Peters, *Biochim. Biophys. Acta*, 1998, **1378**, F115-77.
144. M. Olivier, R. Eeles, M. Hollstein, M. A. Khan, C. C. Harris, and P. Hainaut, *Hum. Mutat.*, 2002, **19**, 607–614.
145. P. A. Jones, W. M. Rideout, J.-C. Shen, C. H. Spruck, and Y. C. Tsai, *BioEssays*, 1992, **14**, 33–36.
146. K. Ory, Y. Legros, C. Auguin, and T. Soussi, *EMBO J.*, 1994, **13**, 3496–504.
147. A. N. Bullock, J. Henckel, and A. R. Fersht, *Oncogene*, 2000, **19**, 1245–1256.
148. L. R. Dearth, H. Qian, T. Wang, T. E. Baroni, J. Zeng, S. W. Chen, S. Y. Yi, and R. K. Brachmann, *Carcinogenesis*, 2006, **28**, 289–298.
149. A. C. Joerger, H. C. Ang, D. B. Veprintsev, C. M. Blair, and A. R. Fersht, *J. Biol. Chem.*, 2005, **280**, 16030–16037.
150. A. C. Joerger, H. C. Ang, and A. R. Fersht, *Proc. Natl. Acad. Sci.*, 2006, **103**, 15056–15061.
151. A. Eldar, H. Rozenberg, Y. Diskin-Posner, R. Rohs, and Z. Shakked, *Nucleic Acids Res.*, 2013, **41**, 8748–8759.

152. H. C. Ang, A. C. Joerger, S. Mayer, and A. R. Fersht, *J. Biol. Chem.*, 2006, **281**, 21934–21941.

153. A. C. Joerger, H. C. Ang, and A. R. Fersht, *Proc. Natl. Acad. Sci.*, 2006, **103**, 15056–15061.

154. A. C. Joerger and A. R. Fersht, *Oncogene*, 2007, **26**, 2226–2242.

155. B. D. Wallentine, Y. Wang, V. Tretyachenko-Ladokhina, M. Tan, D. F. Senear, and H. Luecke, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2013, **69**, 2146–2156.

156. O. Suad, H. Rozenberg, R. Brosh, Y. Diskin-Posner, N. Kessler, L. J. W. Shimon, F. Frolow, A. Liran, V. Rotter, and Z. Shakked, *J. Mol. Biol.*, 2009, **385**, 249–265.

157. D. Malkin, F. P. Li, L. C. Strong, J. F. Fraumeni, C. E. Nelson, D. H. Kim, J. Kassel, M. A. Gryka, F. Z. Bischoff, and M. A. Tainsky, *Science*, 1990, **250**, 1233–8.

158. F. P. Li, J. F. Fraumeni, J. J. Mulvihill, W. A. Blattner, M. G. Dreyfus, M. A. Tucker, and R. W. Miller, *Cancer Res.*, 1988, **48**, 5358–62.

159. S. Masciari, A. Dewanwala, E. M. Stoffel, G. Y. Lauwers, H. Zheng, M. I. Achatz, D. Riegert-Johnson, L. Foretova, E. M. Silva, L. Digianni, S. J. Verselis, K. Schneider, F. P. Li, J. Fraumeni, J. E. Garber, and S. Syngal, *Genet. Med.*, 2011, **13**, 651–657.

160. M. W. G. Ruijs, S. Verhoef, M. A. Rookus, R. Pruntel, A. H. van der Hout, F. B. L. Hogervorst, I. Kluijt, R. H. Sijmons, C. M. Aalfs, A. Wagner, M. G. E. M. Ausems, N. Hoogerbrugge, C. J. van Asperen, E. B. Gomez Garcia, H. Meijers-Heijboer, L. P. ten Kate, F. H. Menko, and L. J. van 't Veer, *J. Med. Genet.*, 2010, **47**, 421–428.

161. P. Wong, S. J. Verselis, J. E. Garber, K. Schneider, L. DiGianni, D. H. Stockwell, F. P. Li, and S. Syngal, *Gastroenterology*, 2006, **130**, 73–79.

162. K. E. Nichols, D. Malkin, J. E. Garber, J. F. Fraumeni, and F. P. Li, *Cancer Epidemiol. Biomarkers Prev.*, 2001, **10**, 83–7.

163. M. Hisada, J. E. Garber, C. Y. Fung, J. F. Fraumeni, and F. P. Li, *J. Natl. Cancer Inst.*, 1998, **90**, 606–11.

164. F. Lalloo, J. Varley, D. Ellis, A. Moran, L. O'Dair, P. Pharoah, D. G. R. Evans, and Early Onset Breast Cancer Study Group, *Lancet*, 2003, **361**, 1101–1102.

165. M. I. W. Achatz, P. Hainaut, and P. Ashton-Prolla, *Lancet Oncol.*, 2009, **10**, 920–925.

166. S. Garritano, F. Gemignani, E. I. Palmero, M. Olivier, G. Martel-Planche, F. Le Calvez-Kelm, L. Brugiéres, F. R. Vargas, R. R. Brentani, P. Ashton-Prolla, S. Landi, S. V. Tavtigian, P. Hainaut, and M. I. W. Achatz, *Hum. Mutat.*, 2010, **31**, 143–150.

167. E. I. Palmero, L. Schüler-Faccini, M. Caleffi, M. I. W. Achatz, M. Olivier, G. Martel-Planche, V. Marcel, E. Aguiar, J. Giacomazzi, I. P. Ewald, R. Giugliani, P. Hainaut, and P. Ashton-Prolla, *Cancer Lett.*, 2008, **261**, 21–25.

168. E. L. DiGiammarino, A. S. Lee, C. Cadwell, W. Zhang, B. Bothner, R. C. Ribeiro, G. Zambetti, and R. W. Kriwacki, *Nat. Struct. Biol.*, 2002, **9**, 12–16.

169. D. Malkin, *Genes Cancer*, 2011, **2**, 475–84.

170. S. Kraiss, A. Quaiser, M. Oren, and M. Montenarh, *J. Virol.*, 1988, **62**, 4737–44.

171. S. Kraiss, S. Spiess, E. Reihsaus, and M. Montenarh, *Exp. Cell Res.*, 1991, **192**, 157–64.

172. J. S. Butler and S. N. Loh, *Biochemistry*, 2003, **42**, 2396–2403.

173. D. Ishimaru, L. R. Andrade, L. S. P. Teixeira, P. A. Quesado, L. M. Maiolino, P. M. Lopez, Y. Cordeiro, L. T. Costa, W. M. Heckl, G. Weissmüller, D. Foguel, and J. L. Silva, *Biochemistry*, 2003, **42**, 9022–9027.

174. A. P. D. Ano Bom, L. P. Rangel, D. C. F. Costa, G. A. P. de Oliveira, D. Sanches, C. A. Braga, L. M. Gava, C. H. I. Ramos, A. O. T. Cepeda, A. C. Stumbo, C. V. De Moura Gallo, Y. Cordeiro, and J. L. Silva, *J. Biol. Chem.*, 2012, **287**, 28152–28162.

175. A. N. Bullock and A. R. Fersht, *Nat. Rev. Cancer*, 2001, **1**, 68–76.

176. J. Milner and E. A. Medcalf, *Cell*, 1991, **65**, 765–74.

177. J. Xu, J. Reumers, J. R. Couceiro, F. De Smet, R. Gallardo, S. Rudyak, A. Cornelis, J. Rozenski, A. Zwolinska, J.-C. Marine, D. Lambrechts, Y.-A. Suh, F. Rousseau, and J. Schymkowitz, *Nat. Chem. Biol.*, 2011, **7**, 285–295.

178. G. Wang and A. R. Fersht, *Proc. Natl. Acad. Sci.*, 2015, **112**, 2443–2448.

179. R. Wilcken, X. Liu, M. O. Zimmermann, T. J. Rutherford, A. R. Fersht, A. C. Joerger, and F. M. Boeckler, *J. Am. Chem. Soc.*, 2012, **134**, 6810–8.

180. A. Soragni, D. M. Janzen, L. M. Johnson, A. G. Lindgren, A. T.-Q. Nguyen, E. Tiourin, A. B. Soriaga, J. Lu, L. Jiang, K. F. Faull, M. Pellegrini, S. Memarzadeh, and D. S. Eisenberg, *Cancer Cell*, 2016, **29**, 90–103.

181. G. Lozano and A. J. Levine, *Mol. Carcinog.*, 1991, **4**, 3–9.

182. J. S. Mattick, *Curr. Opin. Genet. Dev.*, 1994, **4**, 823–31.
183. M. M. Candeias, L. Malbert-Colas, D. J. Powell, C. Daskalogianni, M. M. Maslon, N. Naski, K. Bourougaa, F. Calvo, and R. Fåhraeus, *Nat. Cell Biol.*, 2008, **10**, 1098–1105.
184. B. S. Strauss, *Mutat. Res.*, 2000, **457**, 93–104.
185. C. Kimchi-Sarfaty, J. M. Oh, I.-W. Kim, Z. E. Sauna, A. M. Calcagno, S. V. Ambudkar, and M. M. Gottesman, *Science*, 2007, **315**, 525–528.
186. D. Menendez, A. Inga, and M. A. Resnick, *Mol. Cell. Biol.*, 2006, **26**, 2297–2308.
187. Y. Li and C. Prives, *Oncogene*, 2007, **26**, 2220–2225.
188. D. Bergamaschi, M. Gasco, L. Hiller, A. Sullivan, N. Syed, G. Trigiante, I. Yulug, M. Merlano, G. Numico, A. Comino, M. Attard, O. Reelfs, B. Gusterson, A. K. Bell, V. Heath, M. Tavassoli, P. J. Farrell, P. Smith, X. Lu, and T. Crook, *Cancer Cell*, 2003, **3**, 387–402.
189. A. Koushik, R. W. Platt, and E. L. Franco, *Cancer Epidemiol. Biomarkers Prev.*, 2004, **13**, 11–22.
190. S. J. Klug, R. Wilmotte, C. Santos, M. Almonte, R. Herrero, I. Guerrero, E. Caceres, D. Peixoto-Guimaraes, G. Lenoir, P. Hainaut, J. M. Walboomers, and N. Muñoz, *Cancer Epidemiol. Biomarkers Prev.*, 2001, **10**, 1009–12.
191. X. Li, P. Dumont, A. Della Pietra, C. Shetler, and M. E. Murphy, *J. Biol. Chem.*, 2005, **280**, 24245–51.
192. P. Dumont, J. I.-J. Leu, A. C. Della Pietra, D. L. George, and M. Murphy, *Nat. Genet.*, 2003, **33**, 357–365.
193. A. P. Feinberg and B. Vogelstein, *Nature*, 1983, **301**, 89–92.
194. S. B. Baylin and J. E. Ohm, *Nat. Rev. Cancer*, 2006, **6**, 107–116.
195. W. Y. Chen, D. H. Wang, R. C. Yen, J. Luo, W. Gu, and S. B. Baylin, *Cell*, 2005, **123**, 437–448.
196. M. M. Wales, M. A. Biel, W. El Deiry, B. D. Nelkin, J.-P. Issa, W. K. Cavenee, S. J. Kuerbitz, and S. B. Baylin, *Nat. Genet.*, 1995, **1**, 570–577.
197. M. Schroeder and M. J. Mass, *Biochem. Biophys. Res. Commun.*, 1997, **235**, 403–406.
198. M. M. Hossain and S. K. Ray, *J. Cancer Ther.*, 2014, **5**, 1092–1113.
199. D. Jesionek-Kupnicka, M. Szybka, B. Malachowska, W. Fendler, P. Potemski, S. Piaskowski, D. Jaskolski, W. Papierz, W. Skowronski, W. Och, R. Kordek, and I.

Zawlik, *DNA Cell Biol.*, 2014, **33**, 217–26.

200. A. Vilas–Zornoza, X. Agirre, V. Martín-Palanco, J. I. Martín-Subero, E. San José-Eneriz, L. Garate, S. Álvarez, E. Miranda, P. Rodríguez-Otero, J. Rifón, A. Torres, M. J. Calasanz, J. C. Cigudosa, J. Román-Gómez, and F. Prósper, *PLoS One*, 2011, **6**, e17012.
201. I. P. Pogribny and S. J. James, *Cancer Lett.*, 2002, **176**, 169–74.
202. M. Chmelarova, E. Krepinska, J. Spacek, J. Laco, M. Beranek, and V. Palicka, *Clin. Transl. Oncol.*, 2013, **15**, 160–163.
203. J. H. Kang, S. J. Kim, D.-Y. Noh, I. A. Park, K. J. Choe, O. J. Yoo, and H.-S. Kang, *Lab. Investig.*, 2001, **81**, 573–579.
204. E. M. Hurt, S. B. Thomas, B. Peng, and W. L. Farrar, *Cancer Biol. Ther.*, 2006, **5**, 1154–60.
205. P. J. Teoh, T. H. Chung, S. Sebastian, S. N. Choo, J. Yan, S. B. Ng, R. Fonseca, and W. J. Chng, *Leukemia*, 2014, **28**, 2066–74.
206. K. Murao, Y. Kubo, N. Ohtani, E. Hara, and S. Arase, *Br. J. Dermatol.*, 2006, **155**, 999–1005.
207. S. Sidhu, E. Martin, C. Gicquel, J. Melki, S. J. Clark, P. Campbell, C. J. Magarey, K. M. Schulte, H. D. Röher, L. Delbridge, and B. G. Robinson, *Eur. J. Surg. Oncol.*, 2005, **31**, 549–554.
208. P. C. Papaggeli, A. C. Kortsaris, and P. T. Matsouka, *J. BUON.*, **8**, 341–50.
209. R. Radpour, Z. Barekati, M. M. Haghighi, C. Kohler, R. Asadollahi, P. M. Torbati, W. Holzgreve, and X. Y. Zhong, *Mod. Pathol.*, 2010, **23**, 763–772.
210. A. Zhang, X. Pan, Y. Xia, Q. Xiao, and X. Huang, *Zhonghua Yu Fang Yi Xue Za Zhi*, 2011, **45**, 393–8.
211. R. Najjar Sadeghi, M. Vahedi, H. Zojaji, and M. R. Zali, *J. Dig. Dis.*, 2013, **14**, 305–310.
212. H. Zeng, L. Yan, W.-H. Cheng, and E. O. Uthus, *J. Nutr.*, 2011, **141**, 1464–1468.
213. M. Wade, Y.-C. Li, and G. M. Wahl, *Nat. Rev. Cancer*, 2013, **13**, 83–96.
214. M. Krajewski, P. Ozdowy, L. D'Silva, U. Rothweiler, and T. A. Holak, *Nat. Med.*, 2005, **11**, 1135–1136.
215. M. Wanzel, J. B. Vischedyk, M. P. Gittler, N. Gremke, J. R. Seiz, M. Hefter, M. Noack, R. Savai, M. Mernberger, J. P. Charles, J. Schneikert, A. C. Bretz, A. Nist, and T. Stiewe, *Nat. Chem. Biol.*, 2015, **12**, 22–28.

216. N. Issaeva, P. Bozko, M. Enge, M. Protopopova, L. G. G. C. Verhoef, M. Masucci, A. Pramanik, and G. Selivanova, *Nat. Med.*, 2004, **10**, 1321–1328.
217. M. Wade, Y. C. Li, A. S. Matani, S. M. G. Braun, F. Milanesi, L. W. Rodewald, and G. M. Wahl, *Oncogene*, 2012, **31**, 4789–4797.
218. Y. Yang, R. L. Ludwig, J. P. Jensen, S. A. Pierre, M. V. Medaglia, I. V. Davydov, Y. J. Safiran, P. Oberoi, J. H. Kenten, A. C. Phillips, A. M. Weissman, and K. H. Vousden, *Cancer Cell*, 2005, **7**, 547–559.
219. L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi, and E. A. Liu, *Science*, 2004, **303**, 844–848.
220. Y. Rew, D. Sun, X. Yan, H. P. Beck, J. Canon, A. Chen, J. Duquette, J. Eksterowicz, B. M. Fox, J. Fu, A. Z. Gonzalez, J. Houze, X. Huang, M. Jiang, L. Jin, Y. Li, Z. Li, Y. Ling, M.-C. Lo, A. M. Long, L. R. McGee, J. McIntosh, J. D. Oliner, T. Osgood, A. Y. Saiki, P. Shaffer, Y. C. Wang, S. Wortman, P. Yakowec, Q. Ye, D. Yu, X. Zhao, J. Zhou, J. C. Medina, and S. H. Olson, *J. Med. Chem.*, 2014, **57**, 10499–10511.
221. S. A. Showalter, L. Bruschweiler-Li, E. Johnson, F. Zhang, and R. Brüschweiler, *J. Am. Chem. Soc.*, 2008, **130**, 6472–6478.
222. M. Bista, S. Wolf, K. Khoury, K. Kowalska, Y. Huang, E. Wrona, M. Arciniega, G. M. Popowicz, T. A. Holak, and A. Dömling, *Structure*, 2013, **21**, 2143–2151.
223. H. Yang, T. Sanda, E. Lim, A. Burgess, K. Ming Chia, S. Haupt, D. Thomas, and Y. Haupt, *Front. Oncol.*, 2016, **6**.
224. I. Ray-Coquard, J.-Y. Blay, A. Italiano, A. Le Cesne, N. Penel, J. Zhi, F. Heil, R. Rueger, B. Graves, M. Ding, D. Geho, S. A. Middleton, L. T. Vassilev, G. L. Nichols, and B. N. Bui, *Lancet Oncol.*, 2012, **13**, 1133–1140.
225. M. Andreeff, K. R. Kelly, K. Yee, S. Assouline, R. Strair, L. Popplewell, D. Bowen, G. Martinelli, M. W. Drummond, P. Vyas, M. Kirschbaum, S. P. Iyer, V. Ruvolo, G. M. N. Gonzalez, X. Huang, G. Chen, B. Graves, S. Blotner, P. Bridge, L. Jukofsky, S. Middleton, M. Reckner, R. Rueger, J. Zhi, G. Nichols, and K. Kojima, *Clin. Cancer Res.*, 2016, **22**, 868–876.
226. M. H. Aziz, H. Shen, and C. G. Maki, *Oncogene*, 2011, **30**, 4678–4686.
227. S. Francoz, P. Froment, S. Bogaerts, S. De Clercq, M. Maetens, G. Doumont, E. Bellefroid, and J.-C. Marine, *Proc. Natl. Acad. Sci.*, 2006, **103**, 3232–3237.

228. B. Graves, T. Thompson, M. Xia, C. Janson, C. Lukacs, D. Deo, P. Di Lello, D. Fry, C. Garvie, K.-S. Huang, L. Gao, C. Tovar, A. Lovey, J. Wanner, and L. T. Vassilev, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 11788–93.

229. N. Estrada-Ortiz, C. G. Neochoritis, and A. Dömling, *ChemMedChem*, 2016, **11**, 757–772.

230. C. J. Brown, S. T. Quah, J. Jong, A. M. Goh, P. C. Chiam, K. H. Khoo, M. L. Choong, M. A. Lee, L. Yurlova, K. Zolghadr, T. L. Joseph, C. S. Verma, and D. P. Lane, *ACS Chem. Biol.*, 2013, **8**, 506–512.

231. Federico Bernal, Andrew F. Tyler, Stanley J. Korsmeyer, A. Loren D. Walensky, and G. L. Verdine, *J. Am. Chem. Soc.*, 2007, **129**, 2456–2457.

232. Y. S. Chang, B. Graves, V. Guerlavais, C. Tovar, K. Packman, K.-H. To, K. A. Olson, K. Kesavan, P. Gangurde, A. Mukherjee, T. Baker, K. Darlak, C. Elkin, Z. Filipovic, F. Z. Qureshi, H. Cai, P. Berry, E. Feyfant, X. E. Shi, J. Horstick, D. A. Annis, A. M. Manning, N. Fotouhi, H. Nash, L. T. Vassilev, and T. K. Sawyer, *Proc. Natl. Acad. Sci.*, 2013, **110**, E3445–E3454.

233. F. Bernal, M. Wade, M. Godes, T. N. Davis, D. G. Whitehead, A. L. Kung, G. M. Wahl, and L. D. Walensky, *Cancer Cell*, 2010, **18**, 411–422.

234. S. Lain, J. J. Hollick, J. Campbell, O. D. Staples, M. Higgins, M. Aoubala, A. McCarthy, V. Appleyard, K. E. Murray, L. Baker, A. Thompson, J. Mathers, S. J. Holland, M. J. R. Stark, G. Pass, J. Woods, D. P. Lane, and N. J. Westwood, *Cancer Cell*, 2008, **13**, 454–463.

235. H. Vaziri, S. K. Dessain, E. Ng Eaton, S. I. Imai, R. A. Frye, T. K. Pandita, L. Guarente, and R. A. Weinberg, *Cell*, 2001, **107**, 149–59.

236. L. Li, L. Wang, L. Li, Z. Wang, Y. Ho, T. McDonald, T. L. Holyoake, W. Chen, and R. Bhatia, *Cancer Cell*, 2012, **21**, 266–281.

237. X. Lu, O. Ma, T.-A. Nguyen, S. N. Jones, M. Oren, and L. A. Donehower, *Cancer Cell*, 2007, **12**, 342–354.

238. A. G. Gilmartin, T. H. Faitg, M. Richter, A. Groy, M. A. Seefeld, M. G. Darcy, X. Peng, K. Federowicz, J. Yang, S.-Y. Zhang, E. Minthorn, J.-P. Jaworski, M. Schaber, S. Martens, D. E. McNulty, R. H. Sinnamon, H. Zhang, R. B. Kirkpatrick, N. Nevins, G. Cui, B. Pietrak, E. Diaz, A. Jones, M. Brandt, B. Schwartz, D. A. Heerding, and R. Kumar, *Nat. Chem. Biol.*, 2014, **10**, 181–187.

239. B. A. Foster, H. A. Coffey, M. J. Morin, and F. Rastinejad, *Science*, 1999, **286**,

2507–10.

240. T. M. Rippin, V. J. N. Bykov, S. M. V Freund, G. Selivanova, K. G. Wiman, and A. R. Fersht, *Oncogene*, 2002, **21**, 2119–2129.
241. S. Tanner and A. Barberis, *J. Negat. Results Biomed.*, 2004, **3**, 5.
242. V. J. N. Bykov, N. Issaeva, A. Shilov, M. Hultcrantz, E. Pugacheva, P. Chumakov, J. Bergman, K. G. Wiman, and G. Selivanova, *Nat. Med.*, 2002, **8**, 282–288.
243. J. M. R. Lambert, P. Gorzov, D. B. Veprintsev, M. Söderqvist, D. Segerbäck, J. Bergman, A. R. Fersht, P. Hainaut, K. G. Wiman, and V. J. N. Bykov, *Cancer Cell*, 2009, **15**, 376–88.
244. J. Scotcher, D. J. Clarke, S. K. Weidt, C. L. Mackay, T. R. Hupp, P. J. Sadler, and P. R. R. Langridge-Smith, *J. Am. Soc. Mass Spectrom.*, 2011, **22**, 888–897.
245. J. L. Kaar, N. Basse, A. C. Joerger, E. Stephens, T. J. Rutherford, and A. R. Fersht, *Protein Sci.*, 2010, **19**, 2267–2278.
246. V. J. N. Bykov, N. Issaeva, N. Zache, A. Shilov, M. Hultcrantz, J. Bergman, G. Selivanova, and K. G. Wiman, *J. Biol. Chem.*, 2005, **280**, 30384–30391.
247. N. Zache, J. M. R. Lambert, N. Rökaeus, J. Shen, P. Hainaut, J. Bergman, K. G. Wiman, and V. J. N. Bykov, *Mol. Oncol.*, 2008, **2**, 70–80.
248. J. S. Butler and S. N. Loh, *Biochemistry*, 2007, **46**, 2630–2639.
249. X. Yu, A. Vazquez, A. J. Levine, and D. R. Carpizo, *Cancer Cell*, 2012, **21**, 614–625.
250. X. Yu, A. R. Blanden, S. Narayanan, L. Jayakumar, D. Lubin, D. Augeri, S. D. Kimball, S. N. Loh, and D. R. Carpizo, *Oncotarget*, 2014, **5**, 8879–8892.
251. A. R. Blanden, X. Yu, A. J. Wolfe, J. A. Gilleran, D. J. Augeri, R. S. O'Dell, E. C. Olson, S. D. Kimball, T. J. Emge, L. Movileanu, D. R. Carpizo, and S. N. Loh, *Mol. Pharmacol.*, 2015, **87**, 825–831.
252. A. Garufi, D. Trisciuoglio, M. Porru, C. Leonetti, A. Stoppacciaro, V. D'Orazi, M. Avantaggiati, A. Crispini, D. Pucci, and G. D'Orazi, *J. Exp. Clin. Cancer Res.*, 2013, **32**, 72.
253. D. Pucci, T. Bellini, A. Crispini, I. D'Agnano, P. F. Liguori, P. Garcia-Orduña, S. Pirillo, A. Valentini, and G. Zanchetta, *Med. Chem. Commun.*, 2012, **3**, 462.
254. A. C. Joerger, R. Wilcken, and A. Andreeva, *Structure*, 2014, **22**, 1301–10.
255. S. Gordo, V. Martos, E. Santos, M. Menendez, C. Bo, E. Giralt, and J. de Mendoza, *Proc. Natl. Acad. Sci.*, 2008, **105**, 16426–16431.

256. R. Kamada, W. Yoshino, T. Nomura, Y. Chuman, T. Imagawa, T. Suzuki, and K. Sakaguchi, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 4412–4415.
257. A. Okal, S. Cornillie, S. J. Matissek, K. J. Matissek, T. E. Cheatham, and C. S. Lim, *Mol. Pharm.*, 2014, **11**, 2442–2452.
258. A. C. Joerger, M. R. Bauer, R. Wilcken, M. G. J. Baud, H. Harbrecht, T. E. Exner, F. M. Boeckler, J. Spencer, and A. R. Fersht, *Structure*, 2015, **23**, 2246–2255.
259. X. Liu, R. Wilcken, A. C. Joerger, I. S. Chuckowree, J. Amin, J. Spencer, and A. R. Fersht, *Nucleic Acids Res.*, 2013, **41**, 6034–44.
260. R. Wilcken, G. Wang, F. M. Boeckler, and A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 13584–9.
261. F. M. Boeckler, A. C. Joerger, G. Jaggi, T. J. Rutherford, D. B. Veprintsev, and A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 10360–5.
262. M. Visscher, M. R. Arkin, and T. B. Dansen, *Curr. Opin. Chem. Biol.*, 2016, **30**, 61–67.
263. D. J. Jollow, J. R. Mitchell, W. Z. Potter, D. C. Davis, J. R. Gillette, and B. B. Brodie, *J. Pharmacol. Exp. Ther.*, 1973, **187**, 195–202.
264. D. Gunnell, V. Murray, and K. Hawton, *Suicide Life. Threat. Behav.*, 2000, **30**, 313–26.
265. Y. M. K. And and R. P. Hanzlik, *Chem. Res. Toxicol.*, 2002, **15**, 699–706.
266. I. S. Dunn, D. J. Liberato, N. Castagnoli, and V. S. Byers, *Cell. Immunol.*, 1982, **74**, 220–233.
267. V. R. Thompson and A. P. DeCaprio, *Chem. Res. Toxicol.*, 2013, **26**, 1263–1271.
268. I. M. Serafimova, M. A. Pufall, S. Krishnan, K. Duda, M. S. Cohen, R. L. Maglathlin, J. M. McFarland, R. M. Miller, M. Frödin, and J. Taunton, *Nat. Chem. Biol.*, 2012, **8**, 471–6.
269. James C. Powers, Juliana L. Asgian, A. Özlem Doğan Ekici, and K. E. James, *Chem. Rev.*, 2002, **102**, 4639–4750.
270. L. F. Povirk and D. E. Shuker, *Mutat. Res.*, 1994, **318**, 205–26.
271. S. Zhou, E. Chan, W. Duan, M. Huang, and Y.-Z. Chen, *Drug Metab. Rev.*, 2005, **37**, 41–213.
272. D. C. Evans, A. P. Watt, D. A. Nicoll-Griffith, and T. A. Baillie, *Chem. Res. Toxicol.*, 2004, **17**, 3–16.
273. J. Uetrecht, *Chem. Res. Toxicol.*, 2009, **22**, 24–34.

274. S. N. Lavergne, B. K. Park, and D. J. Naisbitt, *Curr. Opin. Allergy Clin. Immunol.*, 2008, **8**, 299–307.
275. J. Singh, R. C. Petter, T. A. Baillie, and A. Whitty, *Nat. Rev. Drug Discov.*, 2011, **10**, 307–317.
276. C. González-Bello, *ChemMedChem*, 2016, **11**, 22–30.
277. C.-H. Yun, K. E. Mengwasser, A. V. Toms, M. S. Woo, H. Greulich, K.-K. Wong, M. Meyerson, and M. J. Eck, *Proc. Natl. Acad. Sci.*, 2008, **105**, 2070–2075.
278. T. M. Penning, *Trends Pharmacol. Sci.*, 1983, **4**, 212–217.
279. C. Walsh, *Tetrahedron*, 1982, **38**, 871–909.
280. R. R. Rando, A. A. Patchett, S. Marbury, A. L. Maycock, L. M. Perkins, G. A. Doldouras, D. E. Duggan, and S. D. Aster, *Trends Pharmacol. Sci.*, 1979, **1**, 168–171.
281. L. Olbe, E. Carlsson, and P. Lindberg, *Nat. Rev. Drug Discov.*, 2003, **2**, 132–139.
282. W. B. Im, J. C. Sih, D. P. Blakeman, and J. P. McGrath, *J. Biol. Chem.*, 1985, **260**, 4591–7.
283. P. Lindberg, P. Nordberg, T. Alminger, A. Braendstroem, and B. Wallmark, *J. Med. Chem.*, 1986, **29**, 1327–1329.
284. J. M. Shin, Y. M. Cho, and G. Sachs, *J. Am. Chem. Soc.*, 2004, **126**, 7800–7811.
285. Anna E. Speers, A. Gregory C. Adam, and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686–4687.
286. A. E. Speers and B. F. Cravatt, *Chem. Biol.*, 2004, **11**, 535–546.
287. C. M. Salisbury and B. F. Cravatt, *QSAR Comb. Sci.*, 2007, **26**, 1229–1238.
288. M. D. Best, *Biochemistry*, 2009, **48**, 6571–6584.
289. J. P. Alexander and B. F. Cravatt, *Chem. Biol.*, 2005, **12**, 1179–1187.
290. K. Ahn, D. S. Johnson, M. Mileni, D. Beidler, J. Z. Long, M. K. McKinney, E. Weerapana, N. Sadagopan, M. Liimatta, S. E. Smith, S. Lazerwith, C. Stiff, S. Kamtekar, K. Bhattacharya, Y. Zhang, S. Swaney, K. Van Becelaere, R. C. Stevens, and B. F. Cravatt, *Chem. Biol.*, 2009, **16**, 411–420.
291. N. Basse, J. L. Kaar, G. Settanni, A. C. Joerger, T. J. Rutherford, and A. R. Fersht, *Chem. Biol.*, 2010, **17**, 46–56.
292. Remsen. Ira and Norris.F.James, *J. Chem. Soc. Abstr.*, 1896, **70**, 336–337.
293. F. Guthrie, *J. Chem. Soc.*, 1863, **16**, 239–244.
294. P. Murray-Rust and W. D. S. Motherwell, *J. Am. Chem. Soc.*, 1979, **101**, 4374–

4376.

295. P. Murray-Rust, W. C. Stallings, C. T. Monti, R. K. Preston, and J. P. Glusker, *J. Am. Chem. Soc.*, 1983, **105**, 3206–3214.
296. T. Clark, M. Hennemann, J. S. Murray, and P. Politzer, *J. Mol. Model.*, 2007, **13**, 291–296.
297. P. Politzer, J. S. Murray, and M. C. Concha, *J. Mol. Model.*, 2007, **13**, 643–650.
298. P. Metrangolo, J. S. Murray, T. Pilati, P. Politzer, G. Resnati, and G. Terraneo, *Cryst. Growth Des.*, 2011, **11**, 4238–4246.
299. D. Chopra and T. N. G. Row, *CrystEngComm*, 2011, **13**, 2175.
300. P. Politzer, J. S. Murray, and T. Clark, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7748.
301. P. Politzer, P. Lane, M. C. Concha, Y. Ma, and J. S. Murray, *J. Mol. Model.*, 2007, **13**, 305–311.
302. Y. Lu, T. Shi, Y. Wang, H. Yang, X. Yan, X. Luo, H. Jiang, and W. Zhu, *J. Med. Chem.*, 2009, **52**, 2854–2862.
303. K. E. Riley, J. S. Murray, J. Fanfrlík, J. Řezáč, R. J. Solá, M. C. Concha, F. M. Ramos, and P. Politzer, *J. Mol. Model.*, 2011, **17**, 3309–3318.
304. K. E. Riley, J. S. Murray, J. Fanfrlík, J. Řezáč, R. J. Solá, M. C. Concha, F. M. Ramos, and P. Politzer, *J. Mol. Model.*, 2013, **19**, 4651–4659.
305. P. Auffinger, F. A. Hays, E. Westhof, and P. S. Ho, *Proc. Natl. Acad. Sci.*, 2004, **101**, 16789–16794.
306. C. Ouvrard, J.-Y. Le Questel, M. Berthelot, and C. Laurence, *Acta Crystallogr. Sect. B Struct. Sci.*, 2003, **59**, 512–526.
307. R. Wilcken, M. O. Zimmermann, A. Lange, A. C. Joerger, and F. M. Boeckler, *J. Med. Chem.*, 2013, **56**, 1363–1388.
308. Y. Lu, Y. Wang, W. Zhu, P. Hobza, T. N. G. Row, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, R. J. Fletterick, T. E. Steinbacher, W. A. Schumacher, N. Grazier, A. Pudzianowski, A. Apedo, L. Discenza, J. Yanchunas, P. D. Stein, and K. S. Atwal, *Phys. Chem. Chem. Phys.*, 2010, **12**, 4543.
309. R. Wilcken, M. O. Zimmermann, A. Lange, S. Zahn, B. Kirchner, and F. M. Boeckler, *J. Chem. Theory Comput.*, 2011, **7**, 2307–2315.
310. M. Alam, R. E. Beevers, T. Ceska, R. J. Davenport, K. M. Dickson, M. Fortunato, L. Gowers, A. F. Haughan, L. A. James, M. W. Jones, N. Kinsella, C. Lowe, J. W.

G. Meissner, A.-L. Nicolas, B. G. Perry, D. J. Phillips, W. R. Pitt, A. Platt, A. J. Ratcliffe, A. Sharpe, and L. J. Tait, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 3463–3467.

311. J. S. Murray, P. Lane, and P. Politzer, *J. Mol. Model.*, 2009, **15**, 723–729.
312. J. S. Murray, P. Lane, T. Clark, and P. Politzer, *J. Mol. Model.*, 2007, **13**, 1033–1038.
313. C. A. Morgado, J. P. McNamara, I. H. Hillier, N. A. Burton, and M. A. Vincent, *J. Chem. Theory Comput.*, 2007, **3**, 1656–1664.
314. D. Pal and P. Chakrabarti, *J. Biomol. Struct. Dyn.*, 2001, **19**, 115–128.
315. S. H. Gellman, *Biochemistry*, 1991, **30**, 6633–6636.
316. G. Duan, V. H. Smith Jr, and D. F. Weaver, *Mol. Phys.*, 2001, **99**, 1689–1699.
317. R. J. Zauhar, C. L. Colbert, R. S. Morgan, and W. J. Welsh, *Biopolymers*, 2000, **53**, 233–248.
318. A. R. Viguera and L. Serrano, *Biochemistry*, 1995, **34**, 8771–8779.
319. C. D. Tatko and M. L. Waters, *Protein Sci.*, 2004, **13**, 2515–2522.
320. G. D. Probst, S. Bowers, J. M. Sealy, A. P. Truong, R. K. Hom, R. A. Galemmo, A. W. Konradi, H. L. Sham, D. A. Quincy, H. Pan, N. Yao, M. Lin, G. Tóth, D. R. Artis, W. Zmolek, K. Wong, A. Qin, C. Lorentzen, D. F. Nakamura, K. P. Quinn, J.-M. Sauer, K. Powell, L. Ruslim, S. Wright, D. Chereau, Z. Ren, J. P. Anderson, F. Bard, T. A. Yednock, and I. Griswold-Prenner, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 315–319.
321. S. Purser, P. R. Moore, S. Swallow, and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320–30.
322. K. Muller, C. Faeh, and F. Diederich, *Science*, 2007, **317**, 1881–1886.
323. P. Zhou, J. Zou, F. Tian, and Z. Shang, *J. Chem. Inf. Model.*, 2009, **49**, 2344–2355.
324. R. Paulini, K. Müller, and F. Diederich, *Angew. Chemie Int. Ed.*, 2005, **44**, 1788–1805.
325. E. S. Manas, R. J. Unwalla, Z. B. Xu, M. S. Malamas, C. P. Miller, H. A. Harris, C. Hsiao, T. Akopian, W.-T. Hum, K. Malakian, S. Wolfrom, A. Bapat, R. A. Bhat, M. L. Stahl, W. S. Somers, and J. C. Alvarez, *J. Am. Chem. Soc.*, 2004, **126**, 15106–15119.
326. C. Bissantz, B. Kuhn, and M. Stahl, *J. Med. Chem.*, 2010, **53**, 5061–5084.
327. S. W. Cowan-Jacob, G. Fendrich, A. Floersheimer, P. Furet, J. Liebetanz, G.

Rummel, P. Rheinberger, M. Centeleghe, D. Fabbro, and P. W. Manley, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2007, **63**, 80–93.

328. C. D. Cox, M. J. Breslin, B. J. Mariano, P. J. Coleman, C. A. Buser, E. S. Walsh, K. Hamilton, H. E. Huber, N. E. Kohl, M. Torrent, Y. Yan, L. C. Kuo, and G. D. Hartman, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 2041–2045.
329. C. Graebe, C. Glaser, *Berichte der Dtsch. Chem. Gesellschaft*, 1872, **5**, 12.
330. D. P. Chakraborty, B. K. Barman, and P. K. Bose, *Tetrahedron*, 1965, **21**, 681–685.
331. K. C. Das, D. P. Chakraborty, and P. K. Bose, *Experientia*, 1965, **21**, 340–340.
332. P. K. Chakraborty, D.P., Barman, B.K., Boser, *Sci. Cult.*, 1964, **30**, 445.
333. A. D. Favia, D. Habrant, R. Scarpelli, M. Migliore, C. Albani, S. M. Bertozzi, M. Dionisi, G. Tarozzo, D. Piomelli, A. Cavalli, and M. De Vivo, *J. Med. Chem.*, 2012, **55**, 8807–8826.
334. V. Di Bussolo and F. Minutolo, *ChemMedChem*, 2011, **6**, 2133–2136.
335. N. Noguchi, K. Nishino, and E. Niki, *Biochem. Pharmacol.*, 2000, **59**, 1069–1076.
336. H. De Jesús-Cortés, P. Xu, J. Drawbridge, S. J. Estill, P. Huntington, S. Tran, J. Britt, R. Tesla, L. Morlock, J. Naidoo, L. M. Melito, G. Wang, N. S. Williams, J. M. Ready, S. L. McKnight, and A. A. Pieper, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 17010–5.
337. X. Wu, J. Kosaraju, W. Zhou, and K. Y. Tam, *ACS Chem. Neurosci.*, 2017, **8**, 676–685.
338. N. Campbell and B. M. Barclay, *Chem. Rev.*, 1947, **40**, 359–380.
339. M. Majchrzak, M. Grzelak, and B. Marciniec, *Org. Biomol. Chem.*, 2016, **14**.
340. J. B. Kyziol and Z. Daszkiewicz, *Tetrahedron*, 1984, **40**, 1857–1861.
341. A. Cipiciani, S. Clementi, P. Linda, G. Marino, and G. Savelli, *J. Chem. Soc., Perkin Trans. 2*, 1977, 1284–1287.
342. R. Xiong and K. Borbas, *Synlett*, 2014, **26**, 484–488.
343. A. R. Katritzky, G. W. Rewcastle, and L. M. Vazquez de Miguel, *J. Org. Chem.*, 1988, **53**, 794–799.
344. A. Cipiciani, S. Clementi, P. Linda, G. Marino, and G. Savelli, *J. Chem. Soc., Perkin Trans. 2*, 1977, 1284–1287.
345. B. Robinson, *Chem. Rev.*, 1963, **63**, 373–401.
346. B. Robinson, *Chem. Rev.*, 1969, **69**, 227–250.

347. W. Borsche, *Justus Liebig's Ann. der Chemie*, 1908, **359**, 49–80.
348. E. Fischer and F. Jourdan, *Berichte der Dtsch. Chem. Gesellschaft*, 1883, **16**, 2241–2245.
349. B.-Y. Lim, M.-K. Choi, and C.-G. Cho, *Tetrahedron Lett.*, 2011, **52**, 6015–6017.
350. F. R. Japp and F. Klingemann, *Berichte der Dtsch. Chem. Gesellschaft*, 1887, **20**, 2942–2944.
351. F. R. Japp and F. Klingemann, *Berichte der Dtsch. Chem. Gesellschaft*, 1887, **20**, 3284–3286.
352. F. R. Japp and F. Klingemann, *Berichte der Dtsch. Chem. Gesellschaft*, 1887, **20**, 3398–3401.
353. F. R. Japp and F. Klingemann, *Justus Liebig's Ann. der Chemie*, 1888, **247**, 190–225.
354. C. Graebe and F. Ullmann, *Justus Liebig's Ann. der Chemie*, 1896, **291**, 16–17.
355. F. Ullmann, *Justus Liebig's Ann. der Chemie*, 1904, **332**, 82–104.
356. R. W. G. Preston, S. H. Tucker, and J. M. L. Cameron, *J. Chem. Soc.*, 1942, 500–504.
357. J. I. G. Cadogan, M. Cameron-Wood, R. K. Mackie, and R. J. G. Searle, *J. Chem. Soc.*, 1965, 4831.
358. B. Iddon, O. Meth-Cohn, E. F. V. Scriven, H. Suschitzky, and P. T. Gallagher, *Angew. Chemie Int. Ed. English*, 1979, **18**, 900–917.
359. B. G. Soderberg, *Curr. Org. Chem.*, 2000, **4**, 727–764.
360. J. H. Smitrovich and I. W. Davies, *Org. Lett.*, 2004, **6**, 533–535.
361. A. W. Freeman, M. Urvoy, and M. E. Criswell, *J. Org. Chem.*, 2005, **35**, 5014–5019.
362. A. de la Hoz, Á. Díaz-Ortiz, and A. Moreno, *Chem. Soc. Rev.*, 2005, **34**, 164–178.
363. P. Lidström, J. Tierney, B. Wathey, and J. Westman, *Tetrahedron*, 2001, **57**, 9225–9283.
364. Y. Qiu, D. Ma, C. Fu, and S. Ma, *Org. Biomol. Chem.*, 2013, **11**, 1666–71.
365. J. K. Kwon, J. H. Cho, Y.-S. Ryu, S. H. Oh, and E. K. Yum, *Tetrahedron*, 2011, **67**, 4820–4825.
366. W. Gurova, Katerina. Wade, 2014.
367. S. Bhattacharyya, *J. Org. Chem.*, 1995, **60**, 4928–4929.
368. W. Gurova, Katerina. Wade, 2014.

369. A. F. Abdel-Magid and S. J. Mehrman, *Org. Process Res. Dev.*, 2006, **10**, 971–1031.
370. P. N. Riley, P. E. Fanwick, I. P. Rothwell, and T. Ch, *Chem. Commun.*, 1997, **9504**, 1109–1110.
371. A. Vulpetti and C. Dalvit, *Drug Discov. Today*, 2012, **17**, 890–897.
372. H.-J. Böhm, D. Banner, S. Bendels, M. Kansy, B. Kuhn, K. Müller, U. Obst-Sander, and M. Stahl, *ChemBioChem*, 2004, **5**, 637–643.
373. W. K. Hagmann, *J. Med. Chem.*, 2008, **51**, 4359–4369.
374. S. Purser, P. R. Moore, S. Swallow, and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320–330.
375. E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly, and N. A. Meanwell, *J. Med. Chem.*, 2015, **58**, 8315–8359.
376. M. Morgenthaler, E. Schweizer, A. Hoffmann-Röder, F. Benini, R. E. Martin, G. Jaeschke, B. Wagner, H. Fischer, S. Bendels, D. Zimmerli, J. Schneider, F. Diederich, M. Kansy, and K. Müller, *ChemMedChem*, 2007, **2**, 1100–1115.
377. L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, D. W. Banner, W. Haap, and F. Diederich, *Angew. Chemie Int. Ed.*, 2011, **50**, 314–318.
378. C. D. Cox, M. J. Breslin, B. J. Mariano, P. J. Coleman, C. A. Buser, E. S. Walsh, K. Hamilton, H. E. Huber, N. E. Kohl, M. Torrent, Y. Yan, L. C. Kuo, and G. D. Hartman, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 2041–2045.
379. D. T. Manallack, *SAR QSAR Environ. Res.*, 2009, **20**, 611–655.
380. A. Avdeef, *Curr. Top. Med. Chem.*, 2001, **1**, 277–351.
381. N. Anderson and J. Borlak, *FEBS Lett.*, 2006, **580**, 5533–5540.
382. B. Fermini and A. A. Fossa, *Nat. Rev. Drug Discov.*, 2003, **2**, 439–447.
383. D. Alberati, D. Hainzl, S. Jolidon, E. A. Krafft, A. Kurt, A. Maier, E. Pinard, A. W. Thomas, and D. Zimmerli, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4311–4315.
384. C. Jamieson, Elizabeth M. Moir, A. Zoran Rankovic, and G. Wishart, *J. Med.*, 2006, **49**, 5029–5046.
385. A. Bondi, *J. Phys. Chem.*, 1964, **68**, 441–451.
386. M. Schlosser and D. Michel, *Tetrahedron*, 1996, **52**, 99–108.
387. Koichi Mikami, A. Yoshimitsu Itoh, and M. Yamanaka, *Chem. Rev.*, 2004, **104**, 1–16.

388. C. Wolf, W. A. König, and C. Roussel, *Liebigs Ann.*, 1995, **1995**, 781–786.
389. F. Leroux, *ChemBioChem*, 2004, **5**, 644–649.
390. M. Zanda, S. Fustero, J. Piera, G. Chiva, M. S. Rosello, M. Sani, M. Zanda, F. Formaggio, M. Zanda, M. Meli, C. R. de Arellano, M. Zanda, R. Sarabu, J. Sepinwall, G. P. Vincent, M. E. Voss, J. Wilken, and D. W. Young, *New J. Chem.*, 2004, **28**, 1401.
391. M. Molteni, C. Pesenti, M. Sani, A. Volonterio, and M. Zanda, *J. Fluor. Chem.*, 2004, **125**, 1735–1743.
392. L. Pauling, *J. Am. Chem. Soc.*, 1932, **54**, 3570–3582.
393. D. O'Hagan, R. L. Abel, M. D. Drake, D. J. Fox, A. K. Ignatz, D. M. Kwiat, K. M. Schaal, P. R. Virkler, T. S. Keith, D. O'Hagan, J. Schneider, F. Diederich, M. Kansy, and K. Müller, *Chem. Soc. Rev.*, 2008, **37**, 308–319.
394. M. Tandon, M.-M. O'Donnell, A. Porte, D. Vensel, D. Yang, R. Palma, A. Beresford, and M. A. Ashwell, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 1709–1712.
395. C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, *Adv. Drug Deliv. Rev.*, 1997, **23**, 3–25.
396. L. Di and E. H. Kerns, in *Drug-Like Properties*, Elsevier, 2016, pp. 51–59.
397. M. Schlosser, *Angew. Chemie Int. Ed.*, 1998, **37**, 1496–1513.
398. M. Sani, A. Volonterio, and M. Zanda, *ChemMedChem*, 2007, **2**, 1693–1700.
399. B. E. Smart, *J. Fluor. Chem.*, 2001, **109**, 3–11.
400. D. Kim, L. Wang, M. Beconi, G. J. Eiermann, M. H. Fisher, H. He, G. J. Hickey, J. E. Kowalchick, B. Leiting, K. Lyons, F. Marsilio, M. E. McCann, R. A. Patel, A. Petrov, G. Scapin, S. B. Patel, R. S. Roy, J. K. Wu, M. J. Wyvratt, B. B. Zhang, L. Zhu, N. A. Thornberry, and A. E. Weber, *J. Med. Chem.*, 2005, **48**, 141–151.
401. M. L. Quan, P. Y. S. Lam, Q. Han, D. J. P. Pinto, M. Y. He, R. Li, C. D. Ellis, C. G. Clark, C. A. Teleha, J.-H. Sun, R. S. Alexander, S. Bai, J. M. Luettgen, R. M. Knabb, P. C. Wong, and R. R. Wexler, *J. Med. Chem.*, 2005, **48**, 1729–1744.
402. D. Y. Buissonneaud, T. van Mourik, and D. O'Hagan, *Tetrahedron*, 2010, **66**, 2196–2202.
403. J. . Edward, *Chem. Ind.*, 1955, 1102.
404. F. Leroux, A. Peter Jeschke, and Manfred Schlosser, *Chem. Rev.*, 2005, **105**, 827–856.
405. E. Juaristi and G. Cuevas, *Tetrahedron*, 1992, **48**, 5019–5087.

406. M. A. Massa, D. P. Spangler, R. C. Durley, B. S. Hickory, D. T. Connolly, B. J. Witherbee, M. E. Smith, and J. A. Sikorski, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 1625–1628.
407. J. D. Dunitz and R. Taylor, *Chem. A Eur. Journal.*, 1997, **3**, 89.
408. C. Bissantz, B. Kuhn, and M. Stahl, *J. Med. Chem*, 2010, **53**, 5061–5084.
409. in *IUPAC Compendium of Chemical Terminology*, IUPAC, Research Triagle Park, NC, 1997.
410. S. Clarke, *J. Chem. Educ.*, 1981, **58**, A246.
411. W. Kauzmann, *Adv. Protein Chem.*, 1959, **14**, 1–63.
412. M. Charton and B. I. Charton, *J. Theor. Biol.*, 1982, **99**, 629–644.
413. L. Englert, A. Biela, M. Zayed, A. Heine, D. Hangauer, and G. Klebe, *Biochim. Biophys. Acta - Gen. Subj.*, 2010, **1800**, 1192–1202.
414. A. P. K. Richard J. Bingham, John B. C. Findlay, Shih-Yang Hsieh, S. E. V. P. Alexandra Kjellberg, Chiara Perazzolo, W. B. T. Kothandaraman Seshadri, Chi H. Trinh, and and S. W. H. Geoffrey Bodenhausen, *J. Am. Chem. Soc.*, 2004, **126**, 1675–1681.
415. S. W. Homans, *Drug Discov. Today*, 2007, **12**, 534–539.
416. A. Biela, F. Sielaff, F. Terwesten, A. Heine, T. Steinmetzer, and G. Klebe, *J. Med. Chem.*, 2012, **55**, 6094–6110.
417. K. J. Breslauer, D. P. Remeta, W. Y. Chou, R. Ferrante, J. Curry, D. Zaunczkowski, J. G. Snyder, and L. A. Marky, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 8922–6.
418. D. B. Smithrud, T. B. Wyman, and F. Diederich, *J. Am. Chem. Soc.*, 1991, **113**, 5420–5426.
419. L. F. Scatena, M. G. Brown, and G. L. Richmond, *Science*, 2001, **292**.
420. J. L. Finney, A. K. Soper, M. Nardone, and J. M. Goodfellow, *Chem. Soc. Rev.*, 1994, **23**, 1.
421. E. A. Meyer, R. K. Castellano, and F. Diederich, *Angew. Chemie Int. Ed.*, 2003, **42**, 1210–1250.
422. N. Shimokhina, A. Bronowska, and S. W. Homans, *Angew. Chemie Int. Ed.*, 2006, **45**, 6374–6376.
423. R. G. Buttery, J. L. Bomben, D. G. Guadagni, and L. C. Ling, *J. Agr.Food.Cehm*, 1971, **19**, 1045.

424. M. C. Chervenak and E. J. Toone, *J. Am. Chem. Soc.*, 1994, **116**, 10533–10539.
425. A. H. Daranas, H. Shimizu, and S. W. Homans, *J. Am. Chem. Soc.*, 2004, **126**, 11870–11876.
426. R. Malham, S. Johnstone, R. J. Bingham, E. Barratt, S. E. V. Phillips, A. Charles A. Laughton, and Steve W. Homans, *J. Am. Chem. Soc.*, 2005, **127**, 17061–17067.
427. T. Young, R. Abel, B. Kim, B. J. Berne, and R. A. Friesner, *Proc. Natl. Acad. Sci.*, 2007, **104**, 808–813.
428. M. Boehringer, H. Fischer, M. Hennig, D. Hunziker, J. Huwyler, B. Kuhn, B. M. Loeffler, T. Luebbers, P. Mattei, R. Narquizian, E. Sebokova, U. Sprecher, and H. P. Wessel, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 1106–1108.
429. J. D. Dunitz, *ChemBioChem*, 2004, **5**, 614–621.
430. C. Dalvit and A. Vulpetti, *ChemMedChem*, 2012, **7**, 262–272.
431. H.-J. Schneider, M. Karplus, G. Dolgonos, J. Lipkowski, H. J. Schneider, B. Schulz, J. G. Vinter, S. Kirpekar, O. L. Malkina, and H. H. Limbach, *Chem. Sci.*, 2012, **3**, 1381.
432. P. Champagne, J. Desroches, and J.-F. Paquin, *Synthesis (Stuttg).*, 2014, **47**, 306–322.
433. C. Dalvit, C. Invernizzi, and A. Vulpetti, *Chem. A Eur. Journal.*, 2014, **20**, 11058–11068.
434. J.-A. van den B. And and K. R. Seddon, *Cryst. Growth Des.*, 2003, **3**, 643–661.
435. E. Carosati, S. Sciabola, and G. Cruciani, *J. Med. Chem.*, 2004, **47**, 5114–5125.
436. E. D'Oria and J. J. Novoa, *CrystEngComm*, 2008, **10**, 423.
437. H. F. Bettinger, *ChemPhysChem*, 2005, **6**, 1169–1174.
438. P. G. Wenthold and R. R. Squires, *J. Phys. Chem.*, 1995, **99**, 2002–2005.
439. C. Laurence, K. A. Brameld, J. Graton, J.-Y. Le Questel, and E. Renault, *J. Med. Chem.*, 2009, **52**, 4073–4086.
440. E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, and D. J. Nesbitt, *Pure Appl. Chem*, 2011, **83**, 1619–1636.
441. E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, and D. J. Nesbitt, *Pure Appl. Chem*, 2011, **83**, 1637–1641.
442. T. Steiner and G. R. Desiraju, *Chem. Commun.*, 1998, 891–892.

443. A. Vulpetti and C. Dalvit, *Drug Discov. Today*, 2012, **17**, 890–897.
444. C. Dalvit, S. Y. Ko, and A. Vulpetti, *J. Fluor. Chem.*, 2013, **152**, 129–135.
445. J. A. Olsen, D. W. Banner, P. Seiler, U. Obst Sander, A. D'Arcy, M. Stihle, K. Müller, and F. Diederich, *Angew. Chemie Int. Ed.*, 2003, **42**, 2507–2511.
446. M. R. Bauer, R. N. Jones, M. G. J. Baud, R. Wilcken, F. M. Boeckler, A. R. Fersht, A. C. Joerger, and J. Spencer, *ACS Chem. Biol.*, 2016, **11**, 2265–2274.
447. M. Suehiro, G. Yang, G. Torchon, E. Ackerstaff, J. Humm, J. Koutcher, and O. Ouerfelli, *Bioorg. Med. Chem.*, 2011, **19**, 2287–97.
448. S. R. Accordino, J. Ariel Rodríguez Fris, and G. A. Appignanesi, *PLoS One*, 2013, **8**, e55123.
449. J. Seco, F. J. Luque, and X. Barril, *J. Med. Chem.*, 2009, **52**, 2363–2371.
450. O. Guvench, A. D. MacKerell, K. Schulten, R. Impey, and M. Klein, *PLoS Comput. Biol.*, 2009, **5**, e1000435.
451. M. Kitayner, H. Rozenberg, N. Kessler, D. Rabinovich, L. Shaulov, T. E. Haran, and Z. Shakked, *Mol. Cell*, 2006, **22**, 741–753.
452. I. Bahar, A. R. Atilgan, and B. Erman, *Fold. Des.*, 1997, **2**, 173–181.
453. I. Bahar, A. Wallqvist, A. D. G. Covell, and R. L. Jernigan, *Biochemistry*, 1998, **37**, 1067–1075.
454. I. Bahar, B. Erman, R. L. Jernigan, A. R. Atilgan, and D. G. Covell, *J. Mol. Biol.*, 1999, **285**, 1023–1037.
455. D. Tobi and I. Bahar, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 18908–13.
456. A. Fernández and M. Lynch, *Nature*, 2011, **474**, 502–505.
457. A. Fernandez, *Transformative Concepts for Drug Design: Target Wrapping*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
458. A. Fernández and L. R. Scott, *Phys. Rev. Lett.*, 2003, **91**, 18102.
459. A. Fernández and R. Scott, *Biophys. J.*, 2003, **85**, 1914–1928.
460. A. Fernández, *Nat. Biotechnol.*, 2004, **22**, 1081–1084.
461. N. Pietrosemoli, A. Crespo, and A. Fernández, *J. Proteome Res.*, 2007, **6**, 3519–3526.
462. E. Schulz, M. Frechero, G. Appignanesi, and A. Fernández, *PLoS One*, 2010, **5**, e12844.
463. S. R. Accordino, J. A. Rodriguez Fris, G. A. Appignanesi, and A. Fernández, *Eur. Phys. J. E*, 2012, **35**, 59.

464. S. R. Accordino, M. A. Morini, M. B. Sierra, J. A. R. Fris, G. A. Appignanesi, and A. Fernández, *Proteins Struct. Funct. Bioinforma.*, 2012, **80**, 1755–1765.
465. S. R. Accordino, J. Ariel Rodríguez Fris, and G. A. Appignanesi, *PLoS Comput. Biol.*, 2013, **8**, e55123.
466. G. Interlandi, S. K. Wetzel, G. Settanni, A. Plückthun, and A. Caflisch, *J. Mol. Biol.*, 2008, **375**, 837–857.
467. J. I. G. Cadogan, M. Cameron-Wood, R. K. Mackie, and R. J. G. Searle, *J. Chem. Soc.*, 1965, 4831.
468. J. H. Smitrovich and I. W. Davies, *Org. Lett.*, 2004, **6**, 533–535.
469. A. Zapf and M. Beller, *Chem. Commun.*, 2005, **0**, 431–440.
470. T. K. Chaitanya and R. Nagarajan, *Org. Biomol. Chem.*, 2011, **9**, 4662–70.
471. S. D. Roughley and A. M. J. Vernalis, *J. Med. Chem*, 2011, **54**, 3451–3479.
472. N. Miyaura and A. Suzuki, *Chem. Rev*, 1995, **95**, 2457–2483.
473. R. Chinchilla and C. Naera, *Chem. Soc. Rev.*, 2011, **40**, 5084–5121.
474. H. H. Hodgson, *Chem. Rev.*, 1947, **40**, 251–277.
475. S. Song, X. Sun, X. Li, Y. Yuan, and N. Jiao, *Org. Lett.*, 2015, **17**, 2886–2889.
476. N. Miyaura, K. Yamada, and A. Suzuki, *Tetrahedron Lett.*, 1979, **20**, 3437–3440.
477. N. Miyaura, K. Yamada, and A. Suzuki, *Tetrahedron Lett.*, 1979, **20**, 3437–3440.
478. A. J. J. Lennox and G. C. Lloyd-Jones, *Chem. Soc. Rev.*, 2014, **43**, 412–443.
479. N. T. S. Phan, M. V. D. Sluys, and C. W. Jones, *Adv. Synth. Catal.*, 2006, **348**, 609–679.
480. A. A. C. Braga, N. H. Morgon, G. U. And, and F. Maseras, *J. Am. Chem. Soc.*, 2005, **127**, 9298–9307.
481. N. Miyaura, *J. Organomet. Chem.*, 2002, **653**, 54–57.
482. Chau Ming So, A. Chak Po Lau, and F. Y. Kwong, *Org. Lett.*, 2007, **9**, 2795–2798.
483. S. Harkal, F. Rataboul, A. Zapf, C. Fuhrmann, T. Riermeier, A. Monsees, and M. Beller, *Adv. Synth. Catal.*, 2004, **346**, 1742–1748.
484. A. F. Littke, C. Dai, and G. C. Fu, *J. Am. Chem. Soc.*, 2000, **122**, 4020–4028.
485. E. Kantchev, C. O'Brien, and M. Organ, *Angew. Chemie Int. Ed.*, 2007, **46**, 2768–2813.
486. U. Christmann and R. Vilar, *Angew. Chemie Int. Ed.*, 2005, **44**, 366–374.
487. A. H. M. de Vries, J. M. C. A. Mulders, J. H. M. Mommers, H. J. W. Henderickx, and J. G. de Vries, *Org. Lett.*, 2003, **5**, 3285–3288.

488. T. Itoh and T. Mase, *Tetrahedron Lett.*, 2005, **46**, 3573–3577.
489. Mark W. Hooper, A. Masaru Utsunomiya, and J. F. Hartwig, *J. Org. Chem.*, 2003, **68**, 2861–2873.
490. I. V. Seregin and V. Gevorgyan, *Chem. Soc. Rev.*, 2007, **36**, 1173.
491. M. A. Düfert, K. L. Billingsley, and S. L. Buchwald, *J. Am. Chem. Soc.*, 2013, **135**, 12877–12885.
492. Q. Shen, S. Shekhar, J. P. Stambuli, and J. F. Hartwig, *Angew. Chemie Int. Ed.*, 2005, **44**, 1371–1375.
493. E. Tyrrell and P. Brookes, *Synthesis (Stuttg).*, 2003, **2003**, 0469–0483.
494. S. D. Walker, T. E. Barder, J. R. Martinelli, and S. L. Buchwald, *Angew. Chemie Int. Ed.*, 2004, **43**, 1871–1876.
495. Anil S. Guram, Anthony O. King, John G. Allen, Xianghong Wang, Laurie B. Schenkel, Johann Chan, Emilio E. Bunel, Margaret M. Faul, Robert D. Larsen, A. Michael J. Martinelli, and P. J. Reider, *Org. Lett.*, 2006, **8**, 1787–1789.
496. J. Tan, Y. Chen, H. Li, and N. Yasuda, *J. Org. Chem.*, 2014, **79**, 8871–8876.
497. C. M. So, C. C. Yeung, C. P. Lau, and F. Y. Kwong, *J. Org. Chem.*, 2008, **73**, 7803–7806.
498. E. Kantchev, C. O'Brien, and M. Organ, *Angew. Chemie Int. Ed.*, 2007, **46**, 2768–2813.
499. W. D. Miller, A. H. Fray, J. T. Quatroche, and C. D. Sturgill, *Org. Process Res. Dev.*, 2007, **11**, 359–364.
500. Z. Ahmadi, J. S. McIndoe, A. Studer, S. Aime, D. Osella, R. Gobetto, L. Milone, M. N. Eberlin, and B. A. D. Neto, *Chem. Commun.*, 2013, **49**, 11488.
501. L. Jedinák, R. Zátopková, H. Zemánková, A. Šustková, and P. Cankař, *J. Org. Chem.*, 2017, **82**, 157–169.
502. K. Tamao, K. Sumitani, and M. Kumada, *J. Am. Chem. Soc.*, 1972, **94**, 4374–4376.
503. R. J. P. Corriu and J. P. Masse, *J. Chem. Soc. Chem. Commun.*, 1972, 144a.
504. M. Yamamura, I. Moritani, and S.-I. Murahashi, *J. Organomet. Chem.*, 1975, **91**, C39–C42.
505. A. de Meijere and F. Diederich, Eds., *Metal-Catalyzed Cross-Coupling Reactions*, Wiley-VCH Verlag GmbH, Weinheim, Germany, 2004.
506. A. M. Berman and J. S. Johnson, *J. Am. Chem. Soc.*, 2004, **126**, 5680–5681.
507. V. P. Mehta and E. V Van Der Eycken, *Chem. Soc. Rev.*, 2011, **40**, 4925–4936.

508. D. Milstein and J. K. Stille, *J. Am. Chem. Soc.*, 1978, **100**, 3636–3638.
509. M. Kosugi, K. Sasazawa, Y. Shimizu, and T. Migita, *Chem. Lett.*, 1977, **6**, 301–302.
510. V. P. Mehta and E. V Van Der Eycken, *Chem. Soc. Rev.*, 2011, **40**, 4925–4936.
511. R. F. Heck and J. P. Nolley, *J. Org. Chem.*, 1972, **37**, 2320–2322.
512. T. Mizoroki, K. Mori, and A. Ozaki, *Bull. Chem. Soc. Jpn.*, 1971, **44**, 581–581.
513. K. Sonogashira, Y. Tohda, and N. Hagihara, *Tetrahedron Lett.*, 1975, **16**, 4467–4470.
514. R. Chinchilla and C. N?jera, *Chem. Soc. Rev.*, 2011, **40**, 5084.
515. L. Xue, Z. Lin, L. Zhu, H. Zhang, A. Lei, H. Zhang, L. Zhu, A. S. Batsanov, J. A. K. Howard, T. B. Marder, and A. Lei, *Chem. Soc. Rev.*, 2010, **39**, 1692–1705.
516. M. Beaupérin, E. Fayad, R. Amardeil, H. Cattey, P. Richard, S. Brandès, P. Meunier, and J.-C. Hierso, *Organometallics*, 2008, **27**, 1506–1513.
517. M. Beaupérin, A. Job, H. Cattey, S. Royer, P. Meunier, and J.-C. Hierso, *Organometallics*, 2010, **29**, 2815–2822.
518. A. Tougerti, S. Negri, and A. Jutand, *Chem. A Eur. Journal.*, 2007, **13**, 666–676.
519. A. S. Guram, R. A. Rennels, and S. L. Buchwald, *Angew. Chemie Int. Ed. English*, 1995, **34**, 1348–1350.
520. J. Louie and J. F. Hartwig, *Tetrahedron Lett.*, 1995, **36**, 3609–3612.
521. J. F. Hartwig, *Acc. Chem. Res.*, 1998, **31**, 852–860.
522. A. W. Freeman, M. Urvoy, and M. E. Criswell, *J. Org. Chem.*, 2005, **70**, 5014–5019.
523. A. Krasovskiy and P. Knochel, *Angew. Chem. Int. Ed. Engl.*, 2004, **43**, 3333–6.
524. J. D. Moseley, P. M. Murray, E. R. Turp, S. N. G. Tyler, and R. T. Burn, *Tetrahedron*, 2012, **68**, 6010–6017.
525. C. Asche and M. Demeunynck, *Anticancer. Agents Med. Chem.*, 2007, **7**, 247–67.
526. L. Di and E. H. Kerns, in *Drug-Like Properties*, Elsevier, 2016, pp. 243–250.
527. M. Fasano, S. Curry, E. Terreno, M. Galliano, G. Fanali, P. Narciso, S. Notari, and P. Ascenzi, *IUBMB Life*, 2005, **57**, 787–796.
528. A. Sharma, K. Singh, and A. Almasan, in *Methods in Molecular Biology*, ed. L. Bjergbaek, Springer New York, Third Edit., 2012, pp. 613–626.
529. C. Asche and M. Demeunynck, *Anticancer. Agents Med. Chem.*, 2007, **7**, 247–67.
530. X. Chen, J. Mihalic, P. Fan, L. Liang, M. Lindstrom, S. Wong, Q. Ye, Y. Fu, J.

Jaen, J.-L. Chen, K. Dai, and L. Li, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 363–6.

531. H. Peng, X. Chen, Y. Chen, Q. He, Y. Xie, and C. Yang, *Tetrahedron*, 2011, **67**, 5725–5731.
532. M. Compain-Batissou, D. Latreche, J. Gentili, N. Walchshofer, and Z. Bouaziz, *Chem. Pharm. Bull.*, 2004, **52**, 1114–1116.
533. J. A. Dodge, T. I. Richardson, C. A. Clarke, S. A. Jones, R. J. Hinklin, C. W. Hummel, and G. S. Lewis, 2010.
534. N. R. Branda, J. G. Finden, S. J. Gauthier, A. Hayek, K. A. Hope-Ross, J. D. Senior, A. Spantulescu, and S. Sviridov, 2012.